NPS ARCHIVE
1997.03
ROZELLE, D.

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA

THESIS

A METHODOLOGY FOR IMPROVING THE USABILITY OF THE ANVIS/HUD COMPUTER BASED TRAINER

by

Daniel R. Rozelle

March 1997

Principal Advisor: Anthony Ciavarelli

Approved for public release; distribution is unlimited

Thesis
R824

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5101

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	March 1997	Master's Thesis

4. TITLE AND SUBTITLE: **A METHODOLOGY FOR IMPROVING THE USABILITY OF THE ANVIS/HUD COMPUTER BASED TRAINER**

5. FUNDING NUMBERS

6. AUTHOR(S) Rozelle, Daniel R.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Postgraduate School
Monterey, CA 93943-5000

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

12a. DISTRIBUTION/AVAILABILITY STATEMENT: Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE: A

13. ABSTRACT *(maximum 200 words)*

Computer software has taken an increasingly larger role in the U.S. Navy. It is used in nearly every facet of naval operations, from administrative chores to controlling complex weapons systems. Because of the high cost of software and the potential for inadvertent misuse, it is important that software be easy to use and understand. This thesis explores the methods and techniques available for conducting software usability evaluations. Using one of the methods described in this thesis, actual software usability testing is done on a recently developed computer-based training (CBT) program. The CBT program evaluated in this study is designed to instruct helicopter pilots in the use of the AN/AVS-7 ANVIS/HUD. The device is an advanced night vision goggle system that is comprised of the AN/AVS-6 ANVIS night vision goggle (NVG) set and a Heads-Up Display (HUD). This thesis describes the usability test conducted on the ANVIS/HUD CBT and establishes a methodology that can be used, not only on future versions of the ANVIS/HUD CBT, but on other PC oriented software. The result of this usability test show that improvement can be made to the navigation method used by the CBT and the presentation of instructional material.

14. SUBJECT TERMS Usability Testing, Computer Based Training, Night Vision Goggles			15. NUMBER OF PAGES **108**
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18 298-102

Approved for public release; distribution is unlimited

A METHODOLOGY FOR IMPROVING THE USABILITY OF THE ANVIS/HUD COMPUTER BASED TRAINER

Daniel R. Rozelle
Lieutenant Commander, United States Navy
B.A., Findlay College, 1985

Submitted in partial fulfillment of the
requirements for the degree of

**MASTER OF SCIENCE IN
INFORMATION TECHNOLOGY MANAGEMENT**

from the

**NAVAL POSTGRADUATE SCHOOL
March 1997**

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5101

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5101

ABSTRACT

Computer software has taken an increasingly larger role in the U.S. Navy. It is used in nearly every facet of naval operations, from administrative chores to controlling complex weapons systems. Because of the high cost of software and the potential for inadvertent misuse, it is important that software be easy to use and understand. This thesis explores the methods and techniques available for conducting software usability evaluations. Using one of the methods described in this thesis, actual software usability testing is done on a recently developed computer-based training (CBT) program. The CBT program evaluated in this study is designed to instruct helicopter pilots in the use of the AN/AVS-7 ANVIS/HUD. The device is an advanced night vision goggle system that is comprised of the AN/AVS-6 ANVIS night vision goggle (NVG) set and a Heads-Up Display (HUD). This thesis describes the usability test conducted on the ANVIS/HUD CBT and establishes a methodology that can be used, not only on future versions of the ANVIS/HUD CBT, but on other PC oriented software. The result of this usability test show that improvement can be made to the navigation method used by the CBT and the presentation of instructional material.

TABLE OF CONTENTS

I. INTRODUCTION 1

A. PURPOSE 1

B. OBJECTIVE 1

C. BACKGROUND 2

1. The AN/AVS-7 ANVIS/HUD 2

2. Computer-Based Training for Night Vision Goggles 2

3. The HH-60H AN/AVS-7 ANVIS/HUD CBT 3

D. PROBLEM STATEMENT 5

E. SUMMARY OF APPROACH 6

F. SCOPE AND LIMITATIONS 7

G. THESIS ORGANIZATION 8

II. SOFTWARE USABILITY AND EVALUATION METHODS 9

A. WHAT IS USABILITY? 9

B. THE PRINCIPLES OF USABILITY 11

C. WHAT IS A USABILITY PROBLEM 12

1. Definition 12

2. Determining Severity 13

D. EVALUATING USABILITY 14

1. Heuristic Evaluation 14

2. Cognitive Walkthrough 15

3. Usability Testing 16

4. Beta Testing 17

E. WHAT METHOD IS BEST FOR EVALUATING THE ANVIS/HUD CBT? 18

1. Usability Testing vs. Beta Testing 20

2. Usability Testing vs. Cognitive Walkthrough 21

3. Usability Testing vs. Heuristic Evaluation 21
4. Conclusion 22
F. CONDUCTING A USABILITY TEST 22
1. Participants 23
2. Deciding What to Measure 23
3. Usability Labs and Field Testing 23
4. Pilot Test 25
5. Having Participants Talk Out Loud During Observation 25
6. Interacting with Participants During Testing 25
7. Avoiding Bias 26
8. Data Tabulation and Analysis 26
III. METHODOLOGY 29
A. SUBJECTS 29
1. Background 29
2. Participants 29
B. INSTRUMENTATION 30
1. Videotaping 30
2. Questionnaire 31
3. Interviews 32
C. PROCEDURES 32
D. ANALYSIS STRATEGY 33
IV. RESULTS AND DISCUSSION 35
A. DATA ANALYSIS 35
B. FINDINGS 35
1. Questionnaire 35
2. Interviews 49
3. Videotapes 49

4. Minimum system requirements 50
C. DISCUSSION 51
V. CONCLUSION AND RECOMMENDATIONS 53
A. CONCLUSION 53
B. RECOMMENDATIONS 53
C. FUTURE RESEARCH 55
APPENDIX A: USABILITY QUESTIONNAIRE FOR THE ANVIS/HUD CBT 57
APPENDIX B: RESEARCH PARTICIPANT CONSENT FORM 63
APPENDIX C: COMMENTS TAKEN FROM QUESTIONNAIRE 67
APPENDIX D: INTERVIEW QUESTIONS AND SIGNIFICANT RESPONSES 69
APPENDIX E: SQUADRON PRESENTATION 73
APPENDIX F: SPREADSHEET DATA FOR THE ANVIS/HUD CBT USABILITY TEST QUESTIONNAIRE 85
LIST OF REFERENCES 89
INITIAL DISTRIBUTION LIST 93

LIST OF FIGURES

Figure 1.1. Block Diagram of the ANVIS/HUD CBT Structure....4

Figure 1.2. Page from a Typical ANVIS/HUD CBT Lesson....5

Figure 2.1. A Model of the Attributes of System Acceptability....10

Figure 2.2. Floor Plan for a Hypothetical, but Typical, Usability Laboratory....24

Figure 2.3. "Triangulating" - Using Multiple Sources of Data to Find the Usability Problems with a Product....27

Figure 4.1. Familiarity with ANVIS/HUD....36

Figure 4.2. Flight Experience with the ANVIS/HUD....36

Figure 4.3. Familiarity with Windows Operating System....37

Figure 4.4. Computer Use by Participants....38

Figure 4.5. Computer Ownership by Participants....38

Figure 4.6. Participants' Computer Application Uses....38

Figure 4.7. Self Description of Participants with Respect to Computers....39

Figure 4.8. Learning to Operate the ANVIS/HUD CBT....39

Figure 4.9. Understanding the ANVIS/HUD CBT Instructions....40

Figure 4.10. Understanding ANVIS/HUD CBT Terminology....40

Figure 4.11. Module Navigation....41

Figure 4.12. Use of VCR Type Controls....42

Figure 4.13. Next Screen in a Sequence....42

Figure 4.14. Going Back to Previous Screen....43

Figure 4.15. Beginning, Middle and End of Tasks....43

Figure 4.16. Order of Instruction 44
Figure 4.17. Order of Screens 44
Figure 4.18. Screen Size 45
Figure 4.19. Text Size 46
Figure 4.20. Graphical Displays 46
Figure 4.21. CBT Diagrams 47
Figure 4.22. Information Arrangement 47
Figure 4.23. Time Spent Using CBT 48

LIST OF TABLES

Table 2.1. Number and Percent of Usability Problems Uncovered by Each Method 19

Table 2.2. Total Problems Found by Evaluation Technique 20

Table 2.3. Number of Problems Found by Technique and Severity 20

Table 2.4. Mean Problem Severity by Technique 20

Table 4.1. Response Statistics for Question 1.1 36

Table 4.2. Response Statistics for Question 2.1 37

Table 4.3. Response Statistics for Question 3.1 40

Table 4.4. Response Statistics for Question 3.2 40

Table 4.5. Response Statistics for Question 3.3 41

Table 4.6. Response Statistics for Question 4.1 42

Table 4.7. Response Statistics for Question 4.2 42

Table 4.8. Response Statistics for Question 4.3 43

Table 4.9. Response Statistics for Question 4.4 43

Table 4.10. Response Statistics for Question 4.5 44

Table 4.11. Response Statistics for Question 4.5 44

Table 4.12. Response Statistics for Question 4.7 45

Table 4.13. Response Statistics for Question 5.1 45

Table 4.14. Response Statistics for Question 5.2 46

Table 4.15. Response Statistics for Question 5.3 47

Table 4.16. Response Statistics for Question 5.4 47

Table 4.17. Response Statistics for Question 5.5 48

I. INTRODUCTION

A. PURPOSE

It is the intent of this thesis to develop and apply a methodology that can be used to assess the usability of the current and future versions of the AN/AVS-7 Aviators Night Vision Imaging System/Heads-Up Display (ANVIS/HUD) Computer Based Trainer (CBT). As the Navy, through the efforts of the Naval Aviation Safety School and the Naval Postgraduate School (NPS), moves toward actual use of the ANVIS/HUD software, it is important that the product not only meet instructional requirements, but those for ease of use as well. We must not fail to recognize the needs of the military aviator to have an intuitive, easy to use tool for learning. Otherwise, the risk of investing in, creating and relying on a method of instruction that may be considered too difficult, time consuming or distracting is created.

B. OBJECTIVE

In order to achieve the stated purpose of this thesis three specific objectives were accomplished. First, a selected review of the related literature on software usability evaluation was conducted. This literature review yielded an understanding of the most effective and insightful methods of usability evaluation. Second, using the knowledge from the literature review, a methodology was developed for usability testing and applied in the assessment of the ANVIS/HUD CBT. In addition to the formal investigation of usability, input from helicopter aviators who have used the ANVIS/HUD CBT was solicited. Finally, the results of the usability evaluation were analyzed and summarized in order to capture the benefits of this study for future ANVIS/HUD CBT development.

C. BACKGROUND

1. The AN/AVS-7 ANVIS/HUD

The AN/AVS-7 ANVIS/HUD is an advanced electro-optical system that serves as an aid to rotary-wing pilots and copilots for night flying. This system is a combination of the AN/AVS-6 ANVIS Night Vision Goggle (NVG) set and a Heads-Up Display (HUD). It will ultimately be installed in most of the Navy's 1553 data-bus configured helicopters. The ANVIS is an electro-optical image intensifier system designed to provide aviators with the optimum capability to see in the dark and perform nap-of-the-earth and other terrain flight modes during starlight conditions. With the integrated HUD on the AN/AVS-7 ANVIS/HUD, the pilot and copilot can observe aircraft flight data without having to frequently look down at the flight instrument panel. Information from onboard sensors is displayed as symbology that is superimposed within the ANVIS night vision goggle field of view. The availability of HUD symbology, which depicts such flight and system information as altitude, heading, airspeed, velocity vector, engine status and warning signals, provides essential information from the pilot's normal instrument scan. This helps to reduce crew fatigue and allows the pilot and copilot to maintain an aggressive out-of-cockpit visual scan for improved flight monitoring and situation awareness.

2. Computer Based Training for Night Vision Goggles

In 1992, out of concern for the number of accidents involving NVG flight, Naval Aviation School faculty from NPS participated in a comprehensive study of night vision device training. Results of the study showed that improvements could be made in night vision device training in the areas of improved standard operating procedures, revisions to training doctrine and enhancements to training equipment and simulation devices (Ciavarelli, Sengupta and Baer, 1994). To that end, the Naval Aviation Safety School is in an agreement with Naval Air Systems

Command (NAVAIRSYSCOM) PMA-205 (Training Systems) to develop a CBT application for the AN/AVS-7 ANVIS/HUD unit for naval rotary-wing aircraft. A number of studies have now been completed on the use of CBTs and multimedia to train pilots in the use of NVGs. Most recently the ANVIS/HUD, in pursuit of that goal. Initial work (Bryant and Day, 1994) provided a prototype interactive CBT that incorporated sound, video and animation. Concurrently, Ciavarelli et al., (1994) conducted a study covering NVG training requirements and a review of the technologies available for improving NVG training. These works were followed by a closer look at learning theory and the appropriate use of specific media in presenting content to achieve desired learning objectives (Meza, 1995). A fourth study (Epperson, 1995) was completed that detailed the techniques and requirements for implementing animation within the NVG multimedia CBT system. The culmination of this effort to date is the interactive AN/AVS-7 ANVIS/HUD CBT Program for the HH-60H helicopter developed by Kern and Shaffer (1996). This CBT program is a complete re-engineering of a commercially delivered computer-based ANVIS/HUD trainer originally designed for the UH-1N helicopter. By drawing on the earlier works mentioned above and using portions of code, graphics and text from the UH-1N ANVIS/HUD CBT, a thoroughly revised and updated CBT was produced.

3. The HH-60H AN/AVS-7 ANVIS/HUD CBT

The new HH-60H version of the ANVIS/HUD trainer is composed of five modules. The Overview module is an introduction to the CBT and module 1 provides an introduction to the ANVIS/HUD system. Modules 2, 3 and 4 cover training in symbology, operation and maintenance of the ANVIS/HUD, respectively. Each module contains at least three lessons, several exercises and a test to ensure adequate coverage of the material and evaluation of the student. Figures 1.1 and 1.2 show a block diagram of the trainer's structure and a page

from a typical lesson. A typical flow through Module 1 is highlighted in Figure 1.1. Evident in Figure 1.2 is the interface or screen navigation controls that mimic the operation of a common video cassette recorder (VCR).

A sixth module, called a refresher module, was added after completion of the base system (Foster and Price, 1996). This module explains key points of the HH-60H ANVIS/HUD computer-based trainer and incorporates video of night vision imagery. The refresher module is intended for use as a basic introduction for all rotary wing aviators in the use of the ANIVS/HUD, as refresher for those returning to flight status, or to maintain currency in the ANVIS/HUD operations.

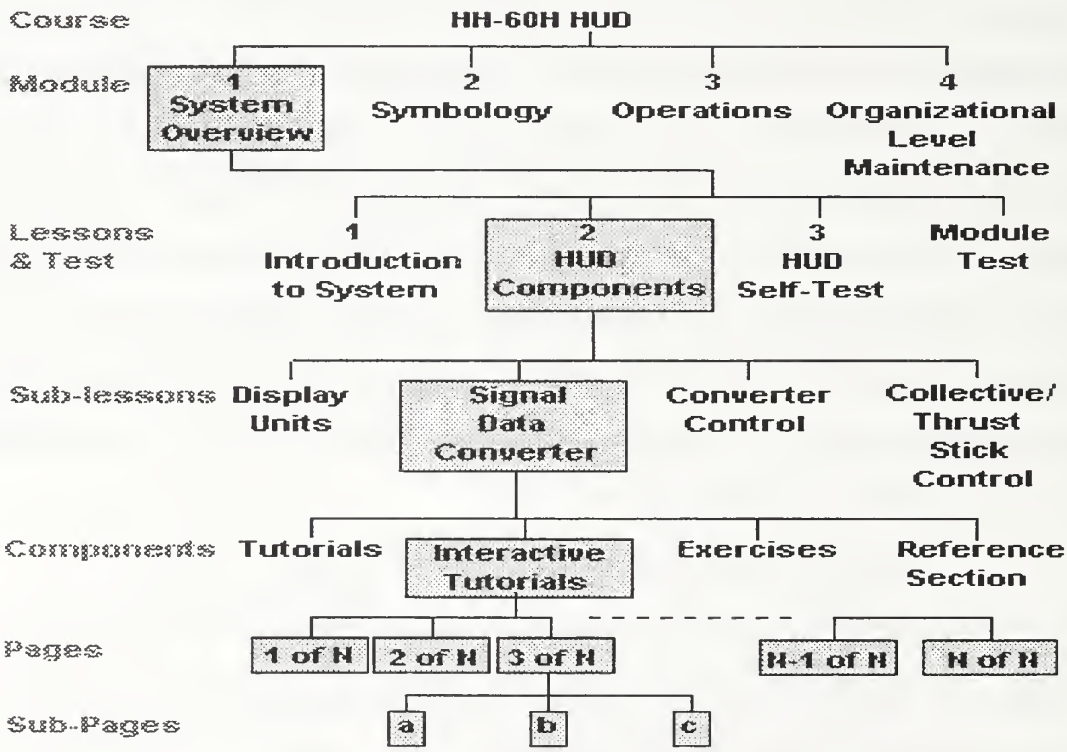


Figure 1.1. Block Diagram of the ANVIS/HUD CBT Structure

(From Kern and Shaffer, 1996).

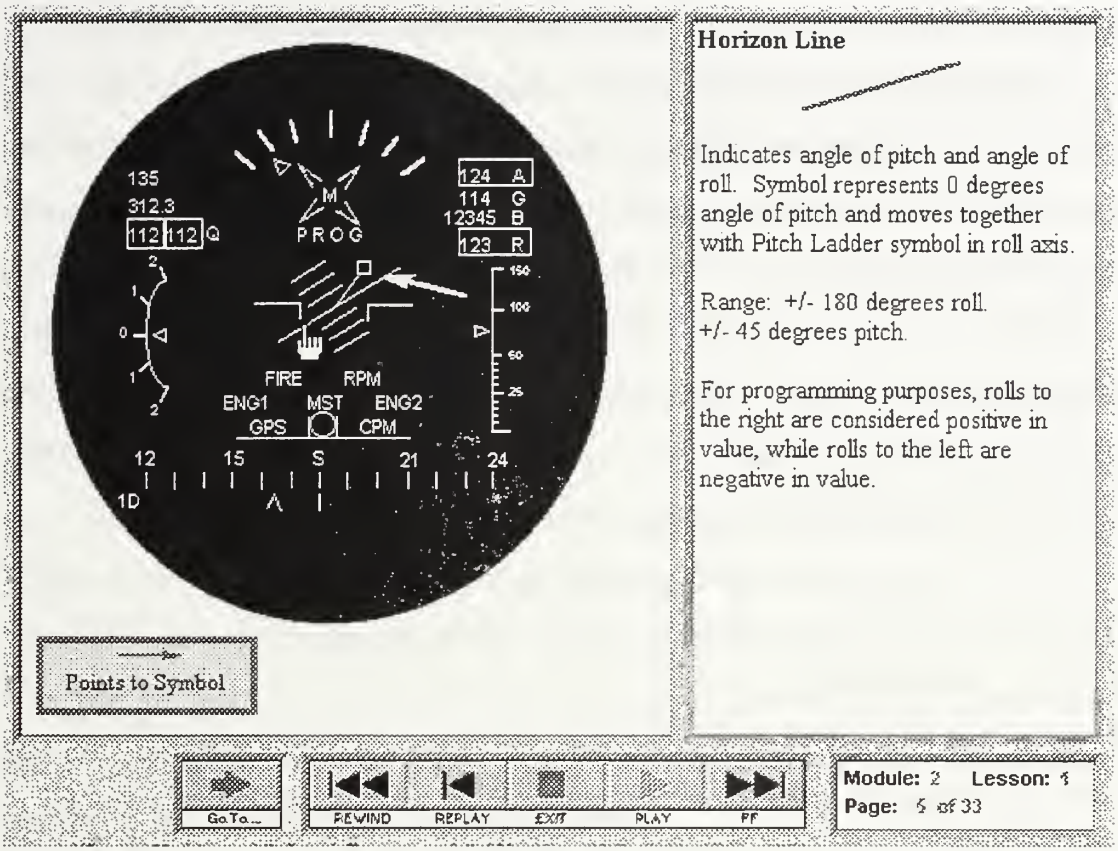


Figure 1.2. Page from a Typical ANVIS/HUD CBT Lesson

(From Kern and Shaffer, 1996).

D. PROBLEM STATEMENT

As previously indicated, at least two years of effort, and over several thousand dollars have been spent in the research and development of the AN/AVS-7 ANVIS/HUD CBT. This effort does not guarantee that a successful learning tool has been produced. No matter how good the instruction content is in the CBT, if the student has difficulty using the computer and interacting with the presentation, then the value of instruction is minimized. Over time, common flaws associated with poor human-computer interface (HCI) design have been identified. Some of the most prevalent problems involve screen presentation formats and navigation between various screens (Kern and Shaffer, 1996). CBT applications often do not effectively describe system functionality to help the user understand

and operate it. Poor design also results from a failure, on the part of the HCI, to reflect the form and level of user guidance considered most appropriate. If the HCI is poorly designed, then the potential exists for the ultimate user of the CBT to reject it simply because it is too difficult or too distracting from which to learn. To avoid this possibility it is important to determine what, if any, ease of use problems may exist in the ANVIS/HUD CBT program's interface design. This determination relies on ascertaining and applying a method of usability evaluation that answers such questions as:

- Are users able to quickly learn how to use the program?
- Is navigation within the program easy to accomplish?
- Are the text based portions of the program easy to read and understand?
- Are the graphics and diagrams easy to see and comprehend?
- Does the arrangement of information on the screen make sense?
- Does the sequence of screens make sense?

By investigating these, and other questions, this study provides a preliminary indication of the success or failure of the ANVIS/HUD CBT and any potential improvements that may be made to the program.

E. SUMMARY OF APPROACH

The general approach to evaluating the usability of the ANVIS/HUD CBT program was to allow HH-60H aviators to interact with the system, capturing their reactions during and immediately after the experience.

The Naval Aviation Safety School provided the ANVIS/HUD CBT program developed by Kern and Shaffer (1996) along with video recording equipment and administrative support. The ANVIS/HUD CBT software was installed on the personal computer hardware located at the HH-60H squadrons. A video camera was then placed near the personal computer in a position that could best capture

both the subjects' reactions while they worked with the CBT and a view of the CBT screen. The videotaped sessions were then examined to highlight any specific usability problems that might not otherwise be discovered.

Immediately following the CBT training sessions, a questionnaire (Appendix A) was filled out by each participant. The questionnaire covered items specific to program navigation, appearance and on-screen instructions. Questions concerning the subject's ANVIS/HUD experience and computer literacy were also included. The questionnaire takes approximately five minutes to complete.

Finally, interviews were conducted with each participant. The questions and significant responses are detailed in Appendix D.

Data collected from the questionnaire was used to determine general areas of the ANVIS/HUD CBT that required changes in presentation or structure to improve usability. The videotaped sessions and interviews provided assistance in pinpointing specific areas needing improvement.

F. SCOPE AND LIMITATIONS

It is not the intent of this research to provide an in-depth study of night vision goggles, night vision goggle training, learning theory or CBT development and use. Prior studies, specifically Kern and Shaffer (1996) and Ciavarelli, et. al, (1994), contain substantial coverage of these subjects. Furthermore, the limitations and usability problems with the current computer hardware needed to run the ANVIS/HUD CBT is well documented in these prior publications and will not be discussed here. The focus of this study is the ANVIS/HUD CBT program or more specifically, the interface provided by the software as it is run on a typical personal computer. Additional points that should be noted when assessing the scope and limitations of this project include:

- The number of subjects available for the study is small. Only two squadrons fly the HH-60H. These squadrons consist primarily of reservists who were generally unavailable for participation.

- A usability lab with significant monitoring and recording capability was not constructed or used. All testing was done in the field with only available equipment.
- Development for the actual HH-60H ANVIS/HUD is still in progress. Likewise, all publications related to the system have not been completed. As work continues on the ANVIS/HUD system, inconsistencies may be identified between the CBT and the actual system.

G. THESIS ORGANIZATION

The next section of this research study is a review of selected literature pertinent to software usability and usability testing. Chapter III covers the methodology behind the data collection and Chapter IV, the analysis of the data collected. The remaining chapter, Chapter V, provides the conclusion and recommendations. The questionnaire, interview transcripts and other items significant to this study are attached as appendices.

II. SOFTWARE USABILITY AND EVALUATION METHODS

A. WHAT IS USABILITY?

All products that we consider tools such as hammers, pens, motor vehicles or even computer software have certain attributes such as color or size. How easy that product is to use is also an attribute. One hammer may be easier to grip than another. A ball point pen glides more smoothly over the writing surface than does a quill pen. A motor vehicle with an automatic transmission is easier to drive than one with a manual transmission. And software that makes use of a graphical user interface is generally considered more pleasing to work with than a command line interface. Today, much greater emphasis is being placed on making products easier to use.

It is important, at this point, to distinguish between a product's functionality and usability. When one talks about the functionality of the product it generally refers to what the product can do. Usability on the other hand refers to how easily and quickly people can accomplish the task for which they are using a product (Dumas and Redish, 1993). For example, many people often consider the functionality of a video cassette recorder (VCR) as the ability to record certain television programs at a pre-selected time or to edit previously recorded videos. The usability of that same VCR would concern how easy it is to accomplish pre-selecting the time and program to record or the specific frames to edit on a video. For many people, even with the number of improvements in functionality, a VCR is not considered to be an easily usable product. For software, a major premise of usability is that a poorly designed interface may result in failure of the system to attain its full functional capability. Thus, in order to improve the usefulness and the corresponding marketability of a product many software developers are beginning to focus on usability.

Usability is only a part of a system's overall acceptability and itself has several attributes. Figure 2.1 shows a simple model of system acceptability and components, including usability and the attributes of usability. It should be noted that "Utility" in the figure represents functionality of the system or product.

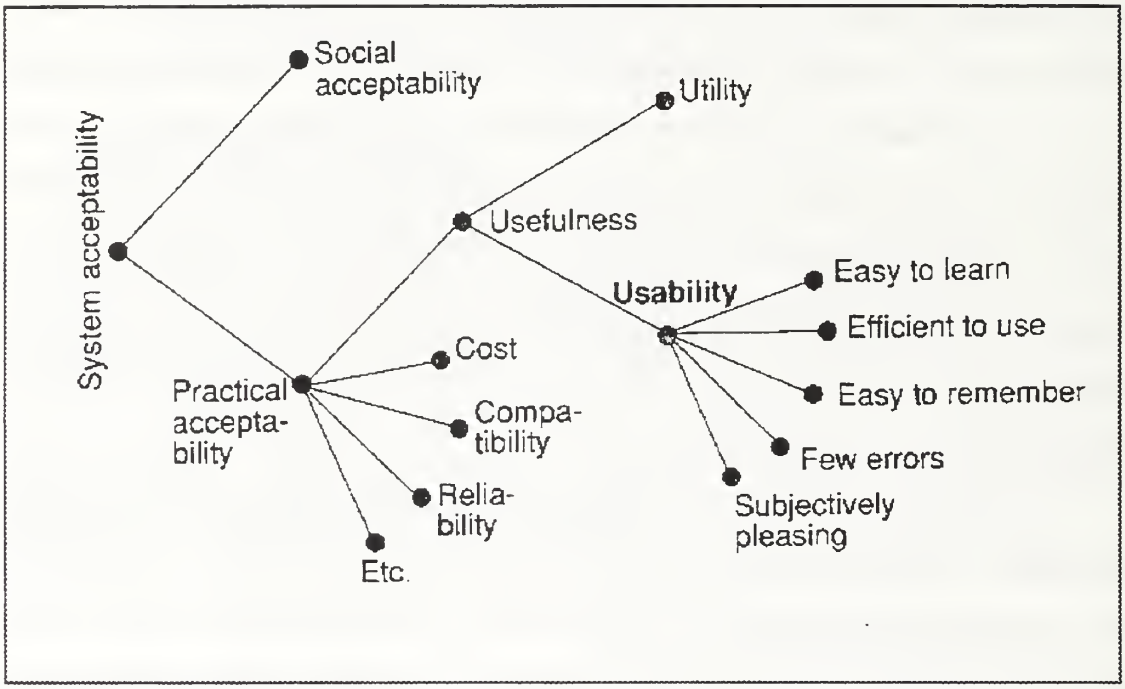


Figure 2.1. A Model of the Attributes of System Acceptability

(From Nielsen, 1993)

When attempting improvement of a product or tool's usability it is important to address each of those components that make up usability.

The process of ensuring a usable software product generally involves two distinct paths, usability *engineering* and usability *evaluation.* Usability engineering occurs in the design phase and encompasses the inclusion of users in product design, adherence to design guidelines for good software, and setting quantitative usability goals such as specific times for task accomplishment. Usability evaluation on the other hand occurs at the end of the design phase. However, it is not dependent on the use of usability engineering in product design.

Instead, usability evaluations are best used as part of a process that focuses on usability throughout design and development, not as the sole time when users are considered (Dumas and Redish, 1993).

B. THE PRINCIPLES OF USABILITY

In an effort to improve software usability and remove the subjectivity from software usability evaluation, many usability experts have developed sets of guidelines or principles of usability. Usability principles, similar to the standards used by software developers for designing applications, help to focus attention on the needs of users. Guidelines, or principles can be both general or specific. A general guideline could be to "provide feedback" to the user. This general advice can be made into a more specific guideline for a graphical user interface: ensure that the main objects of interest to the user are visible on the screen and that their most important attributes are shown (Nielsen, 1993). A few authors in the usability field, most notably Dumas and Redish (1993), differentiate between usability principles as broadly defined goals and usability guidelines as specific measures to achieve usability goals. However, use of this differentiation is inconsistent and the two terms can largely be considered as interchangeable.

Of the numerous different guideline collections that exist, some have so many guidelines that using them is difficult. For example, Smith and Mosier (1986) have developed a set of nearly 1000, Brown (1988) specifies 302 and Mayhew (1992) has 288. However, much of this effort can be reduced to around 15, or fewer, short statements. Dumas and Redish (1993) present several shortened collections including Shneiderman's (1992) "Eight Golden Rules of Dialog Design" which states that interface design should:

1. Strive for consistency
2. Enable frequent users to use shortcuts
3. Offer informative feedback

4. Design dialogs to yield closure
5. Offer simple error handling
6. Permit easy reversal of actions
7. Support internal locus of control, and
8. Reduce short-term memory load

Another list, offered by Karat, Campbell, and Fiegel, (1992), provides a concise compilation of what can generally be considered as the most accepted principles for usability design:

1. Use a simple and natural dialog
2. Provide a intuitive visual layout
3. Speak the user's language
4. Minimize the user's memory load
5. Be consistent
6. Provide feedback
7. Provide clearly marked exits
8. Provide shortcuts
9. Provide good help
10. Allow user customization
11. Minimize the use and effects of modes
12. Support input device continuity

Regardless of generalization, specificity, number or label, usability principles provide an essential basis for creating and evaluating a usable software product.

C. WHAT IS A USABILITY PROBLEM?

1. Definition

When considering software, usability generally refers to the human-computer interface (HCI). Usability attributes of a program's interface may include how one moves around in the program, commonly referred to as the

navigation metaphor, or how easy the text is to read. Other considerations may include consistent and easy to understand wording and appropriate graphics.

Each of the examples of usability attributes mentioned above, and others such as button selection or the colors used, represents a potential usability problem. In broad terms a usability problem can be defined as anything that interferes with a user's ability to efficiently and effectively complete tasks (Karat et al, 1992). More specific to software however, the definition may be stated as: aspects of a user interface that may cause the resulting system to have reduced usability for the end user (Nielsen and Mack, 1994). This directs the focus toward the HCI. A user can find an interface element to be problematic for many reasons. It might make the system harder to learn; it might make it slower for users to perform their tasks; it may cause usage errors; or it may simply be ugly or otherwise unpleasing (Nielsen and Mack, 1994). Because people's tolerance for time spent learning and using tools is low (Dumas and Redish, 1993), attempting to find these problems early in the software development life cycle is crucial to user acceptance.

2. Determining Severity

Determining severity of a usability problem helps to allocate more resources to fix the most serious problems. Such determinations can also provide a rough estimate of the need for additional usability efforts. Severity of a usability problem is largely a combination of three factors:

1. The frequency with which the problem occurs: Is it common or rare?
2. The impact of the problem if it occurs: Will it be easy or difficult for users to overcome?
3. The persistence of the problem: Is it a one time problem that users can overcome once they know about it or will users repeatedly be bothered by the problem?

Even though severity consists of these different components, it is common to combine all aspects of severity in a single severity rating. An overall assessment of each problem can then be made to facilitate prioritizing and decision making. This can be accomplished using ordinal numeric scales to rank the problems. (Nielsen, 1996)

D. EVALUATING USABILITY

Over the past several years three specific methods have evolved for evaluating usability problems in software: heuristic evaluation, cognitive walkthrough and usability testing (often called empirical testing). Heuristic evaluation and cognitive walkthrough are methods in which a software product or design is inspected by individuals or groups familiar with the principles of usability or the software. Usability testing, unlike the other two methods, makes use of real users to actually test the software under controlled conditions. All three methods have advantages and disadvantages that will be more closely examined later in this discussion. Beta testing, while not a true usability evaluation method, is also used as an approach to finding potential problems in software, and as such, is included here in this review of evaluation methods.

1. Heuristic Evaluation

Heuristic evaluation is the most informal of the methods and is conducted simply by looking at an interface and trying to come up with an opinion about what is good and bad about the interface (Nielsen and Molich, 1990). This type of evaluation is conducted by inspecting the software interface for adherence to certain system design rules such as those outlined in the previous section covering usability principles. The principles involved in the evaluation are normally referred to as the heuristics and give the method its name. Generally, it is difficult for a single individual to find all the usability problems in a particular interface. Thus, evaluating an interface through the heuristic inspection method normally

requires between three and five evaluators (Nielsen and Mack, 1994) with each well versed, if not an expert, in the usability principles being applied in the evaluation (Nielsen, 1992).

Heuristic evaluation is performed by having each individual evaluator inspect the interface alone. Only after all evaluations have been completed are the evaluators allowed to communicate and have their findings aggregated. This ensures an independent and unbiased evaluation from each evaluator. Typically, a heuristic evaluation session for an individual evaluator lasts one or two hours. During the evaluation session the evaluator goes through the interface several times, inspecting the various elements and comparing them to the recognized usability principles (Nielsen and Mack, 1994).

2. Cognitive Walkthrough

Originally designed for the evaluation of simple *Walk Up and Use* interfaces (Lewis, Polson, Wharton and Rieman, 1990), the cognitive walkthrough method is meant to be used iteratively and early in a software design cycle. It is a task based method that serves to focus an evaluator's attention on the user's goals and actions, and whether the system supports or hinders the accomplishment of those goals (Wharton, Bradford, Jefferies and Franzke, 1992). Generally software developers or usability specialists perform a walkthrough, either as a group or individually.

During a walkthrough, the steps required to accomplish a task are evaluated by examining how a user would interact with the interface through each step. For each action, the evaluators try to determine what a typical user would be trying to do at each point and what actions the interface allows. In evaluating each step necessary to perform a task, a cognitive walkthrough attempts to uncover design errors that would interfere with learning by exploration. The method finds mismatches between user's and designer's conceptualization of a task, poor

choices of wording for menu titles and button labels, and inadequate feedback about the consequences of an action. The procedure also uncovers implicit and explicit assumptions made by developers about user's knowledge of the task and the interface conventions (Nielsen and Mack, 1994).

A cognitive walkthrough normally consists of three basic phases: a preparation phase, an evaluation phase and a result interpretation phase. Forms, describing the tasks and providing detailed instructions, are used to guide the evaluators through the preparation and evaluation phases. The interpretation phase, on the other hand, is loosely structured. For example, in the preparation phase, the suite of tasks to be evaluated is identified and information about the users is noted. In the evaluation phase, questions are asked concerning each step within a given user task. Finally in the interpretative phase, all information gathered from the walkthrough is interpreted such that negative answers highlight steps or areas that may be difficult for the user (Wharton et al, 1992).

3. Usability Testing

Usability testing with real users is the most fundamental usability evaluation method. It provides direct information about how people use computers and what their exact problems are with the interface being tested (Nielsen, 1993). While there are wide variations in where and how a usability test is conducted, every usability test shares the following five characteristics (Dumas and Redish, 1993):

1. Each test has specific goals to meet the overall primary goal of improving the usability of a product.
2. The participants represent real users.
3. The participants do real tasks.
4. What participants do and say is observed and recorded.
5. The data are analyzed, the real problems are diagnosed and changes are recommended to fix those problems.

Typically, in a usability test of software, specific tasks inherent to the software and a user population for the software are identified. For example, if one was conducting a usability test on a word processing program, then the task might be to find, open and close a file. The population of users may be office secretaries. From the user population several participants are recruited to work with the software in accomplishing the previously identified tasks. During the test the participants' comments concerning the software and their performance while using it are recorded. Finally, the users' opinions about the software are solicited either with a questionnaire or through interviews. It should be noted that what distinguishes a usability test from other data gathering activities such as focus groups, surveys and beta testing, is that the participants' behavior while using the program is always noted or recorded (Dumas and Redish, 1993). In the end a usability test provides both quantitative data using performance measures and qualitative data from the observations and users' comments.

4. Beta Testing

Beta testing is a simple approach to finding problems with software, though not an entirely reliable one. Also known as field testing, clinical trials or user acceptance testing, beta testing is done by making an early version of a program available to users at no cost. Distribution of beta software is usually limited to users who have either volunteered or otherwise been selected to receive the product and participate in the test. No specific tasks are delineated, instead the users, or evaluators, are free to work with the program as they see fit. Ideally, as users use the program they will discover flaws, bugs and usability problems in the completion of their personal tasks. Discovered problems are then reported to the beta test director who in turn ensures that the problems are addressed in subsequent versions of the tested software.

E. WHAT METHOD IS BEST FOR EVALUATING THE ANVIS/HUD CBT?

Several studies have been done that compare the various methods of usability evaluation. Foremost of these studies is the Karat et al (1992) study, the Desurvire, Kondziela and Atwood (1992) study and the Jefferies, Miller, Wharton and Uyeda (1991) study.

Karat et al (1992) conducted a study comparing usability testing to walkthroughs. In this study, two different graphical user interface (GUI) office systems were evaluated to assess the reliability of the methods. In addition to having six individual evaluators conduct a walkthrough, they also had six pairs of evaluators conduct a walkthrough together to see if walkthroughs are made more effective when there is group interaction. The usability tests had six test participants using the product. The results show that the usability tests uncovered about twice as many problems as the group walkthroughs and three times as many as the individual walkthroughs. The usability tests also uncovered significantly more severe problems than the walkthroughs. (Dumas and Redish, 1993)

The Desurvire et al (1992) study compared usability testing against cognitive walkthroughs and heuristic evaluation. Usability experts, software engineers and non-experts were each used for the walkthroughs and heuristic evaluation. The findings from this research are consistent with those of the Karat et al study and are displayed in Table 2.1. As shown, the usability test found more than twice as many problems as the expert heuristic evaluation and more than triple the expert cognitive walkthrough. The usability experts found about twice as many problems as the software engineers and even more than the non-experts (Dumas and Redish, 1993).

Table 2.1. Number and Percent of Usability Problems Uncovered by Each Method

	# of problems	% of Problems
Usability Test	25	100%
Heuristic Evaluation		
Experts	11	44%
Software Engineers	4	16%
Non-Experts	2	8%
Cognitive Walkthrough		
Experts	7	28%
Software Engineers	4	16%
Non-Experts	2	8%

(From Desurvire et al, 1992)

In the Jefferies et al (1991) study the researchers evaluated a pre-release software application using heuristic evaluation, usability testing, guidelines and cognitive walkthrough. Specialized groups were organized around the four techniques to complete the evaluations: Two groups were made up of software engineers. One group conducted a cognitive walkthrough evaluation while the other applied 62 guidelines of good practice in usability as a means of evaluation. For the usability test, a human factors specialist conducted a usability test using six participants to identify usability problems. The final group, made up of user interface (UI) specialists, spent two weeks performing individual heuristic evaluations. Table 2.2 shows the total number of problems found by each evaluation method and the number of core problems (problems not related to the operating system, evaluator error or those not able to be duplicated). Table 2.3 shows the top and bottom thirds of the problems ranked by severity while Table 2.4 displays the mean problem severity by evaluation technique.

Table 2.2. Total Problems Found by Evaluation Technique

	Heuristic Evaluation	Usability Testing	Guidelines	Cognitive Walkthrough	Total
Total	152	38	38	40	268
Core	121	32	35	35	223
Core, no duplicates	105	31	35	35	206

(From Jefferies et al, 1991)

Table 2.3. Number of Problems Found by Technique and Severity

Level of Problem Severity	Heuristic Evaluation	Usability Testing	Guidelines	Cognitive Walkthrough
Most Severe	28	18	12	9
Least Severe	52	2	11	10

(From Jefferies et al, 1991)

Table 2.4. Mean Problem Severity by Technique

Heuristic Evaluation	Usability Testing	Guidelines	Cognitive Walkthrough
3.59	4.15	3.61	3.44

(From Jefferies et al, 1991)

Significant to the results is that none of the heuristic evaluators found more than 42 core problems over the two week period. The usability test however, was completed by the six participants in about five hours (Jefferies et al, 1991).

These three studies allow for summary one-to-one comparisons. For proper inclusion, other sources are used to make the comparison of usability testing to beta testing.

1. Usability Testing vs. Beta Testing

During development of Windows 95, Microsoft Corporation made extensive use of beta testing. Recently however, Microsoft has demonstrated a shift from primarily using beta testing in favor of a more usability oriented

approach (Microsoft, 1997). What advantages does usability testing provide over beta testing? In beta testing, real users do real tasks in their actual work environments. However, many companies find that they get very little feedback from beta testers, and beta testing seldom yields useful information about usability problems for these reasons (Dumas and Redish, 1993):

- The beta test site does not have to use the product.
- Feedback is unsystematic. Users report only what they remember and choose to report.
- No one observes the beta test users and records their behavior.
- Tasks that may be of concern may not even be tested.

The most significant drawback of beta testing though is that it comes too late in the development process. At the time of beta testing the product is usually fully coded and difficult to change.

Usability testing, unlike beta testing can be done throughout the design and development process and users can be observed and recorded as they work with the product or prototypes.

2. Usability Testing vs. Cognitive Walkthrough

Regardless of the use of teams or individuals, cognitive walkthroughs generally tended to fare the poorest, as evidenced by the Karat et al (1992), Desurvire et al (1992) and Jefferies et al (1991) studies. This is because walkthroughs are a technique intended for use by software engineers who do not share the same viewpoint as users. Usability testing, on the other hand, involves real users doing real tasks.

3. Usability Testing vs. Heuristic Evaluation

Heuristic evaluation is the one method that is often used in place of usability testing. In fact a heuristic evaluation can often find more problems than usability testing. (Jefferies et al, 1991). However, it should also be noted that the

heuristic evaluation in the Jefferies et al (1991) study was done over a two week period by four usability interface experts. This is probably an unrealistic period for an evaluation. It is the type of problems found however that distinguishes usability testing from a heuristic evaluation. Usability testing generally finds the more serious and global problem (Dumas and Redish, 1993) while applying heuristics usually uncovers large numbers of specific, one-time (poor grammar or misspellings in a particular text), and low-priority problems (Jefferies et al, 1991). Another disadvantage of heuristic evaluations is a dependence upon having access to several people with the knowledge and experience necessary to apply the technique. Such people are a scarce resource and often multiple evaluators are necessary to obtain significant results (Jefferies et al, 1991). They may be even more difficult to come by if they also need to have expertise in a particular kind of application (Nielsen, 1992).

4. Conclusion

To produce a successful software product it is important to consider the ultimate user and potential problems they may incur while using the software in the everyday completion of their tasks. Without ready availability of usability experts nor the ANVIS/HUD CBT designers, usability testing offers the most effective means of evaluating the ANVIS/HUD CBT software for usability problems and ultimately improving the product.

F. CONDUCTING A USABILITY TEST

Determining the appropriate method for evaluating usability is the first step in a usability evaluation. Once completed the focus turns to actually carrying out the selected method. The use of usability testing for evaluating the ANVIS/HUD CBT requires careful consideration of certain aspects inherent to the method.

1. Participants

Typically, a usability test includes six to 12 participants divided into subgroups. The sub-grouping allows for identification of problems that may be inherent to some characteristic of the participants (Dumas and Redish, 1993). For example, users with significantly more computer experience may have fewer difficulties with a software product being tested. All participants must be members of the group of people who now use or who will use the product.

2. Deciding What To Measure

A usability test should be designed to collect both:

- Performance measures: that is, counts of actions and behaviors that you can see, such as:
  - time to finish a task
  - time spent reading help files
  - number of wrong menu choices
  - observations of frustration
  - observations of confusion
- Subjective measures: that is, people's perceptions, opinions and judgments including:
  - ratings of ease of learning, using or installing a product
  - preferences over previous versions or other products
  - spontaneous comments

Once the desired measures have been chosen then the appropriate testing and collection methods can be designed.

3. Usability Labs and Field Testing

Many usability tests take place in specially equipped usability laboratories. However, usability labs should not be considered an absolute necessity for usability testing. It is possible to accomplish a usability test with no more equipment than a notepad. If a usability lab is desired then nearly any available

office space can be converted into a one. A typical usability lab will have an observation room and a test room separated by sound-proof, one-way mirrors, remote controlled video cameras and video synchronization equipment to mix video of the user and computer screen into a single video stream. Figure 2.2 depicts a floor plan for a typical usability laboratory.

Should a requirement be made that a usability test take place in the field then a portable lab can be devised. For those situations that do require testing outside of a lab, scheduling, test design and equipment setup must be carefully planned (Rowley, 1994).

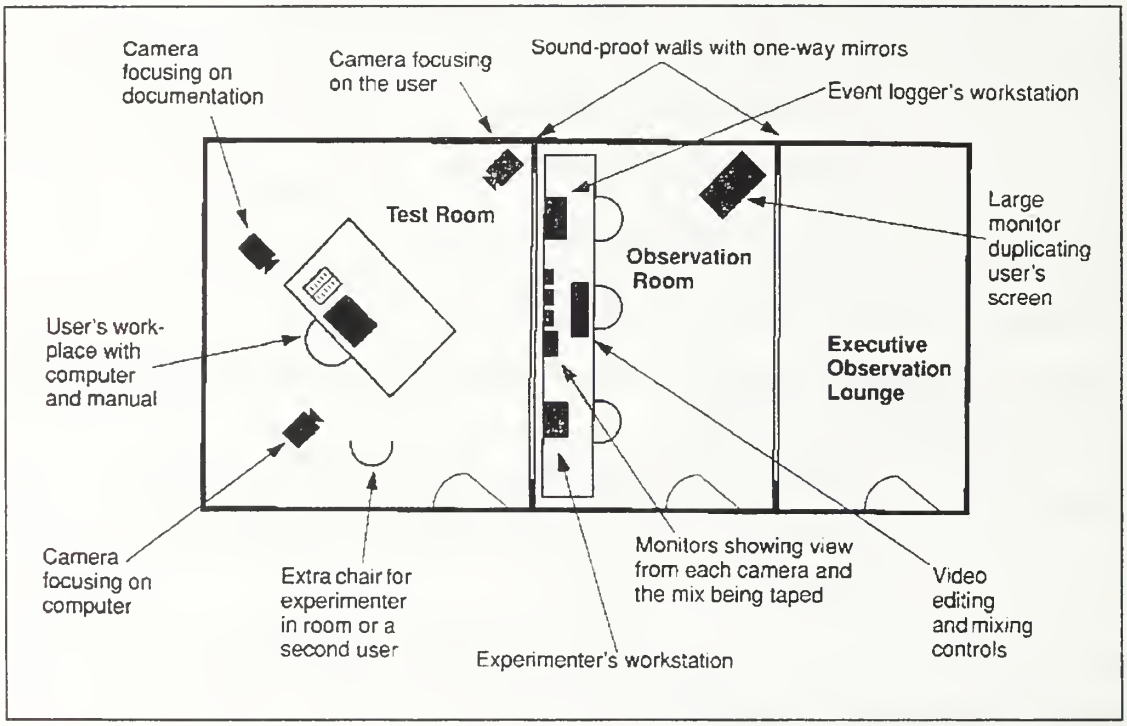

Camera focusing on documentation
Camera focusing on the user
Sound-proof walls with one-way mirrors
Event logger's workstation
Test Room
Observation Room
Large monitor duplicating user's screen
User's work-place with computer and manual
Executive Observation Lounge
Camera focusing on computer
Extra chair for experimenter in room or a second user
Monitors showing view from each camera and the mix being taped
Video editing and mixing controls
Experimenter's workstation

Figure 2.2. Floor Plan for a Hypothetical, but Typical, Usability Laboratory

(From Nielsen, 1993)

4. Pilot Test

No usability testing should be performed without first having tried out the test procedure on pilot subjects. Often, one or two pilot subjects will be enough, but more may be needed for large tests or when the initial pilot tests show severe deficiencies in the test plan (Nielsen, 1993). In the case of the ANVIS/HUD CBT evaluation, the pilot test objective is to "debug" the equipment, materials and procedures. Additionally it provides the opportunity to practice the activities that will be done during the actual usability test. To avoid last minute adjustments the pilot test should be scheduled two days prior to the usability test and should:

- Be conducted exactly as the full usability test
- Use a test participant who represents the targeted users (Dumas and Redish, 1993)

5. Having Participants Talk Out Loud During Observation

For many usability tests, participants are asked to talk or think out loud so that their reactions to the product being tested can be heard. This can help significantly to pinpointing potential problems of products. However, because people do not normally think out loud while they work, they may vary in their ability to express themselves while they work. This in turn leads to two possible drawbacks. First, thinking or talking out loud may cause the participants to take longer in completing tasks, and second it may actually improve participants' performance of the tasks. Neither drawback though, should discourage its use. The value of the information obtained usually outweighs the bias this procedure may cause. (Dumas and Redish, 1993)

6. Interacting With Participants During Testing

During a usability test a number of opportunities may arise that allow interaction between the tester and participants. The goal of the usability test will dictate how interactions are handled. The goal of a test on a near finished product

may be to determine those problems that would cause a user to call for technical support. In that case any interactions such as assistance with the program would be avoided. On the other hand, the test may be on a prototype where the developers desire as much diagnostic information as possible. This situation may lead to continual dialogue between participant and tester. In general, participants should be allowed to do each task at their own pace in their own way (Dumas and Redish, 1993). Regardless of the goal there may be times when certain interactions are necessary. The tester may need to remind the participant to continue talking aloud or to move on to the next task or phase of testing.

7. Avoiding Bias

It is important that interactions with test participants be conducted in a way that avoids introducing bias. Testers can influence the way participants act and what participants say by biasing them with leading questions and encouraging either positive or negative statements. To avoid this possibility caution should be exercised when asking questions of participants or responding to participants' questions and comments (Dumas and Redish, 1993). The interview questions (Appendix D) for the ANVIS/HUD CBT evaluation have been designed to avoid bias using guidance provided by Dumas and Redish (1993).

8. Data Tabulation and Analysis

A usability test generates a substantial amount of data. When the test is complete data may include a list of problems, quantitative data on subjective ratings from the questionnaire, videotapes, participants' comments and user profiles. One technique available for handling such large amounts of data is triangulation (Dumas and Redish, 1993). The concept behind triangulation is to find commonalties among the various collected data which in turn may yield a usability problem. Figure 2.3 is a representation of triangulation.

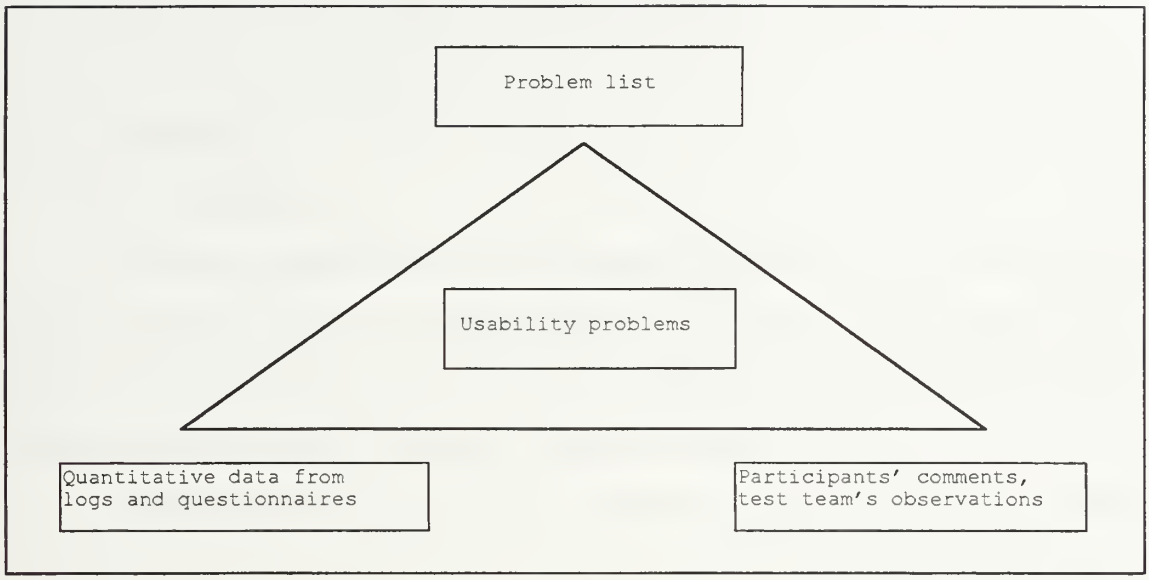


Figure 2.3. "Triangulating" - Using Multiple Sources of Data to Find the Usability Problems With a Product

(From Dumas and Redish, 1994)

Spreadsheet software is normally used in tabulating the data and calculating any statistics required in the analysis. Generally, only simple descriptive statistics and qualitative data, such as test participants' comments, are needed to document the case for the presence of usability problems (Dumas and Redish, 1993).

III. METHODOLOGY

A. SUBJECTS

1. Background

Selection of participants in this research focused on potential users of the ANVIS/HUD CBT. This focus limited the research to the two Helicopter Combat Support Services (HCS) squadrons, HCS-4, located at Norfolk, Virginia and HCS-5, located at Point Mugu, California. Both fly the HH-60H.

The primary missions of an HCS squadron are special warfare and combat search and rescue. Both missions require HCS crews to train for operations in hostile areas, often in the night-time environment. Special warfare involves the insertion and/or extraction of highly specialized teams (i.e., Navy SEALS or Army Rangers) in covert areas. Combat search and rescue involves rescuing downed aviators or stranded service members from a hostile environment or behind enemy lines.

The squadrons consist of roughly 30 percent Training and Administration of Reserves (TAR) personnel and 70 percent Selected Reserves (SELRES) personnel. TARs provide the stability in the squadron by working full work weeks, expediting administrative matters, coordinating deployments, and implementing training plans. On the other hand, in order to cut government costs, selected reserves are required to be at the squadron for only short periods of time (Kern and Shaffer, 1996). As a result of this attendance policy, separating the two groups for evaluation of the CBT was not attempted.

2. Participants

From the available pilots in the squadrons, several were asked to participate in the usability testing of the ANVIS/HUD CBT. All test subjects elected to participate voluntarily. The final sample included seven pilots at HCS-4 and three

at HCS-5. No individuals with specific experience in the use of computers, software or the ANVIS/HUD were singled out for the purpose of including or excluding them from the evaluation. All participants received an overview of the ANVIS/HUD CBT development and an explanation of their role in the usability testing of the trainer. A signed written statement of consent was also obtained from each participant (Appendix B). The participants were asked to be candid in their response to the CBT and were informed that the survey, interview and video data would be confidential. Relevant background data regarding the survey participants is discussed in Chapter IV.

B. INSTRUMENTATION

Evaluation of the usability of the ANVIS/HUD CBT used three data collection methods. This approach was in keeping with the desire to gather both performance and subjective measures as detailed in Chapter II.

1. Videotaping

The Naval Aviation Safety School made available an 8mm video camera. Videotaping the subjects while they use the ANVIS/HUD CBT offered the dual benefit of: 1) capturing their reactions to the software and 2) providing a record for going back and verifying or retracing comments or decisions (Nielsen 1994). During the test the video camera was placed slightly off perpendicular to a line between the subjects and the computer screen. This position provided the best opportunity for capture of both the subjects' reactions while they worked with the CBT and a view of the CBT screen. Subjects were encouraged to talk aloud during their session with the CBT and to point to the screen when encountering what they considered to be a problem with the trainer. Later examination of the videotaped sessions corroborated and provided detail to specific usability problems that might not otherwise be discovered.

2. Questionnaire

A questionnaire (Appendix A) was formulated based on the Questionnaire for User Interface Satisfaction (QUIS) (Chin, Diehl and Norman, 1988; Shneiderman, 1992) and an understanding of potential problems with the ANVIS/HUD CBT.

The questions were organized into six groups. The first and second groups pertained to the participant's familiarity with the ANVIS/HUD and computer experience. The third, fourth and fifth covered learning how to use the CBT, control and navigation of the CBT and appearance issues respectively. The last section consisted of a single question about time spent using the CBT. An open ended comments section was also included so that any concerns not covered by the questionnaire could be addressed. Responses to this section have been recorded and are included in Appendix C.

A Likert type scale was used to indicate strength of opinion on all questions dealing with the usability of the CBT. Descriptive qualifiers relevant to the question asked were used for every answer scale and all scales had a range of 1 to 5 which represented, respectively, low and high rating values. A Not Applicable (NA) option was also included.

A pre-test of the questionnaire was conducted at NPS with both students familiar with the ANVIS/HUD CBT and others who were not. The students not familiar with the ANVIS/HUD CBT were asked to relate the questionnaire to any software program that they had recently used. As a result of this pre-test two specific changes were made:

1. Question 5.1 was modified to eliminate a potential central tendency error.
2. Differentiation was made between HUD graphics and cockpit diagrams

3. Interviews

All participants were interviewed upon completion of the questionnaire. As with the comments section of the questionnaire, the interviews were designed primarily to probe for frustration with specific parts of the CBT. Only four questions were asked:

1. Did you find the portion of the CBT that you worked with to be an easy to use tool for learning?
2. Was there anything that you found particularly difficult to do or understand?
3. If you could change one thing about the CBT to make it easier to use, what would it be?
4. Are there any features you would like to see added to the CBT?

Significant responses to these questions are detailed in Appendix D.

C. PROCEDURES

A pretest walkthrough of equipment setup and instrumentation was completed prior to the field evaluation. The walkthrough took place in the Multimedia Development Lab at the Naval Aviation Safety School and reflected the usability test as planned for the actual testing of the CBT. Participants for this phase were experienced aviators though not necessarily night vision goggle users. During the pretest walkthrough, placement of the video camera was refined and any further ambiguities or typographical errors in the questionnaire or interview questions were corrected.

The first "in-the-field" usability test of the ANVIS/HUD CBT was done at HCS-4 in Norfolk, Virginia. The volunteer participants were given a brief presentation on the ANVIS/HUD CBT program and an explanation of what was desired from them during the test (Appendix E). It was expressly noted to each participant that they were not the subject of the research but were only assisting in the evaluation of the CBT. A statement of informed consent was then obtained

from each participant (Appendix B). After the presentation the subjects were scheduled to use the CBT so as not to interfere with their other daily tasks.

Once the introductory presentation was finished, the CBT and video recording equipment were set up. A relatively secluded area in the squadron offices was obtained to avoid possible interruption during the sessions.

Prior to beginning each individual session the participants were familiarized with the computer hardware and the location of the video equipment. The previous instructions were reiterated and the desire that the participants talk aloud during the session and point to the screen to indicate problems was stated. Participants were then instructed to proceed with the task of working through the introductory and first instructional modules of the CBT. Once the assigned task was finished the participants were given the questionnaire to fill out. Interviews were then done using the video camera to record the responses. Total time to complete all three portions of the test were noted so that a basis for future tests could be established.

D. ANALYSIS STRATEGY

Descriptive statistics were used to examine the data from the questionnaire. The number and percentage of subjects who answer a question in a specific way is shown in a frequency table and illustrated with a bar graph. The number of participants in the test was used as the base for calculating the percentages. The mean, median and standard deviation were also calculated for the answers to each question. All tabulation and calculations were accomplished using Microsoft Excel spreadsheet software.

Significant comments, taken from the open ended portion of the questionnaire and interviews, are quoted in Appendices C and D respectively. Significant events from the videotapes are described in Chapter IV.

IV. RESULTS AND DISCUSSION

A. DATA ANALYSIS

The findings presented in this section have been taken from the post-test questionnaire administered to the ten participants, interviews with the participants and videotapes of the usability tests.

The responses to the questionnaire were transferred to a Microsoft Excel spreadsheet (Appendix F) and simple descriptive statistics were computed from the questions that used Likert scales. The resulting data were converted to tables and figures. Responses to the non-scaled questions were reviewed for any significant differences or similarities among the respondents.

Like the responses to the non-scaled questions of the questionnaire, responses to interviews (Appendix D) were examined for trends indicating problems or frustrations experienced across participants.

Finally the videotapes were viewed so that data from the questionnaire and interviews might be corroborated with specific occurrences. Additionally the videotapes served to highlight any usability problems not revealed by either of the other methods.

B. FINDINGS

1. Questionnaire

The following pages present a detailed statistical and graphical breakdown of responses to each question of the questionnaire. As with the questionnaire, the analysis is divided into distinct subject areas.

a. Knowledge of ANVIS/HUD

Two questions were used to explore participants' experience with the ANVIS/HUD. Most participants indicated little familiarity with the

ANVIS/HUD as shown in Figure 4.1 and Table 4.1. None of the participants in this study had flown with the device (Figure 4.2).

Question 1.1: How familiar are you with the ANVIS/HUD?

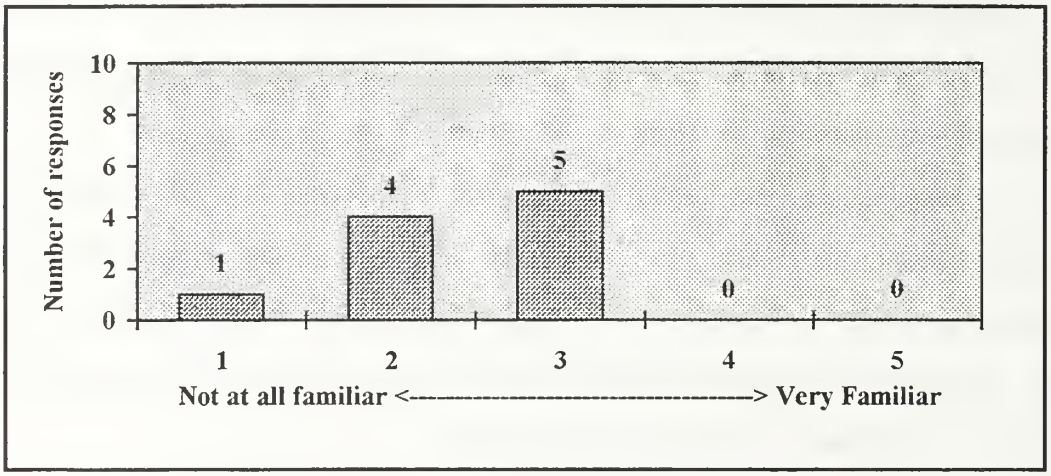


Figure 4.1. Familiarity with ANVIS/HUD

Table 4.1. Response Statistics for Question 1.1

Statistic	Value
Mean	2.40
Median	2.50
Standard Deviation	0.699

Question 1.2: Have you flown with the ANVIS/HUD?

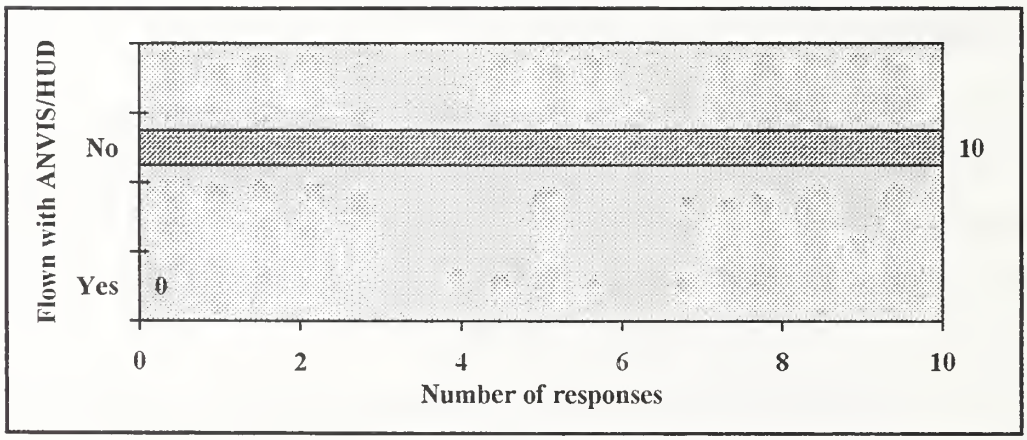


Figure 4.2. Flight Experience with the ANVIS/HUD

b. Participants' Familiarity With Computers

Participants' use and familiarity of computers were examined as a potential indicator of the acceptability of a computer-based trainer. Most considered themselves familiar with the "Windows" operating environment (Figure 4.3 and Table 4.2) and said that they use a computer either frequently or constantly (Figure 4.4). All respondents said that they owned a home computer (Figure 4.5). The most popular use of a computer by participants was for word processing followed by online/Internet uses and databases (Figure 4.6). Finally, when asked to assess themselves regarding computers, eight out of the ten participants identified themselves as regular computer users (Figure 4.7).

Question 2.1: Are you familiar with the "Windows" type operating system?

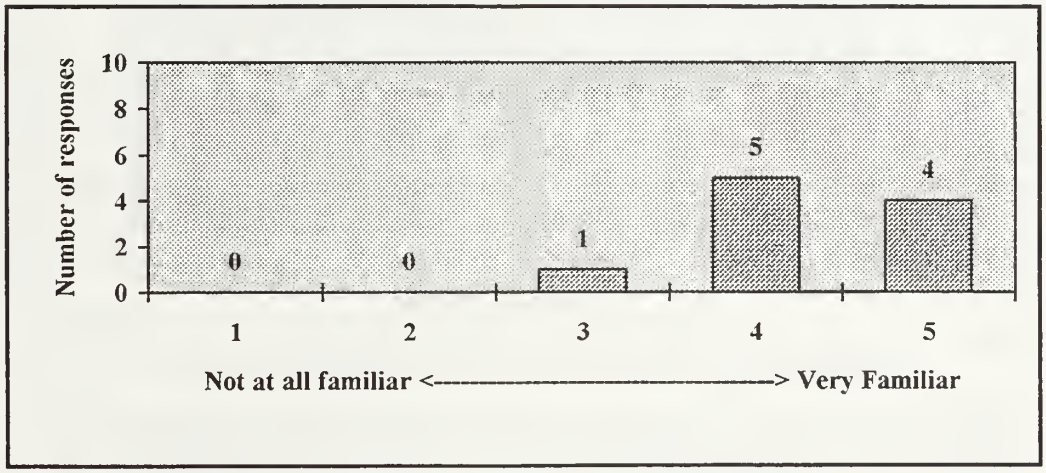


Figure 4.3. Familiarity with Windows Operating System

Table 4.2. Response Statistics for Question 2.1

Statistic	Value
Mean	4.30
Median	4.00
Standard Deviation	0.675

Question 2.2: How often do you use a computer?

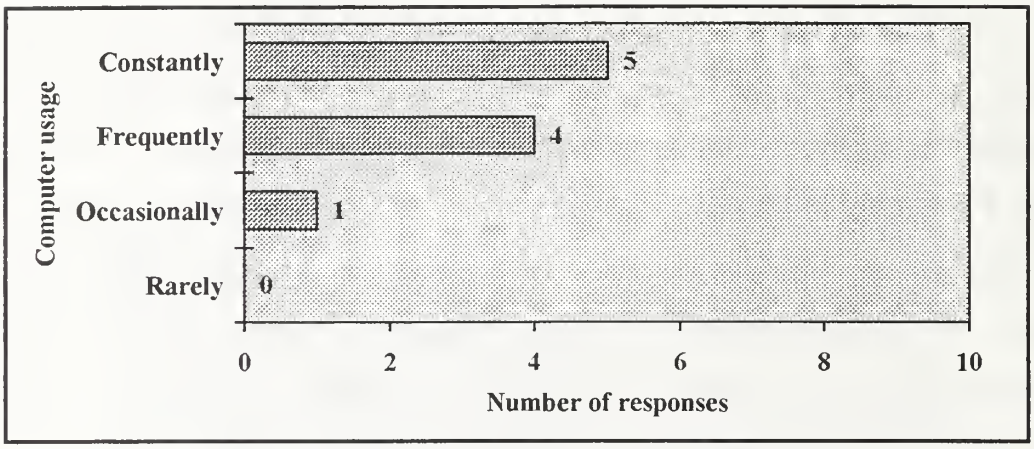


Figure 4.4. Computer Use by Participants

Question 2.3: Do you own a home computer?

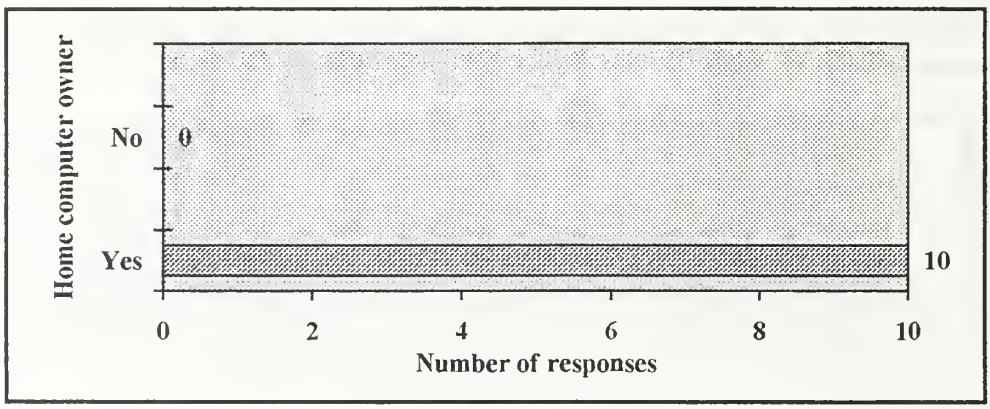


Figure 4.5. Computer Ownership by Participants

Question 2.4: What activities do you most often use the computer for? (circle all that apply)

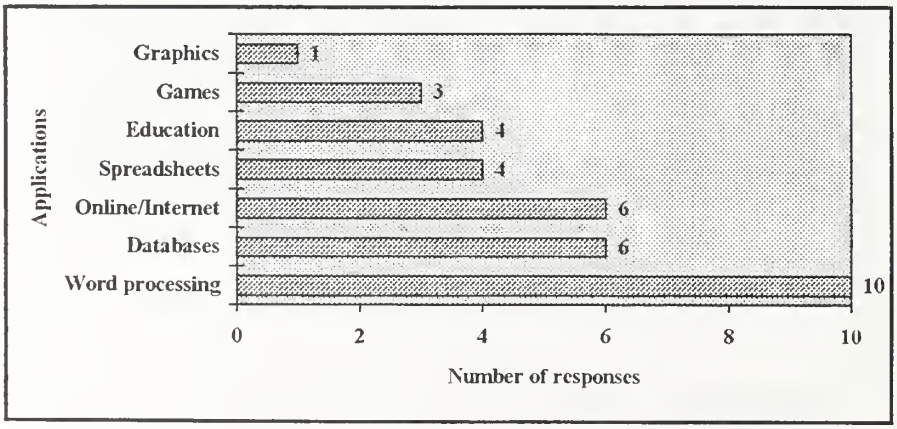


Figure 4.6. Participants' Computer Application Uses

Question 2.5: How would you describe yourself with respect to computers?

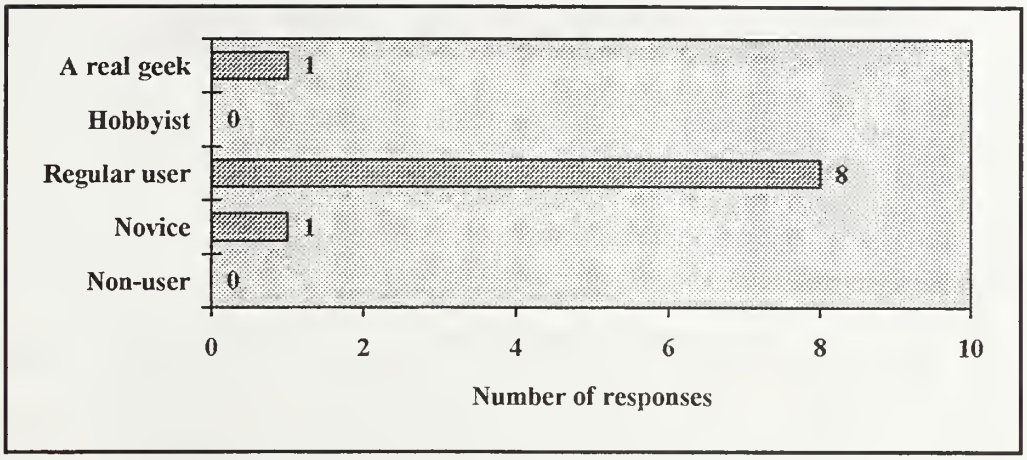


Figure 4.7. Self Description of Participants with Respect to Computers

c. Getting started with the CBT

This group of questions explored the participants experience with the CBT overview and operation instructions. Most felt that the ANVIS/HUD CBT was moderately easy to operate (Figure 4.8 and Table 4.3). Understanding the instructions however, was not considered entirely easy (Figure 4.9 and Table 4.4) though the terminology was generally clear (Figure 4.10 and Table 4.5).

Question 3.1: Learning to operate the program?

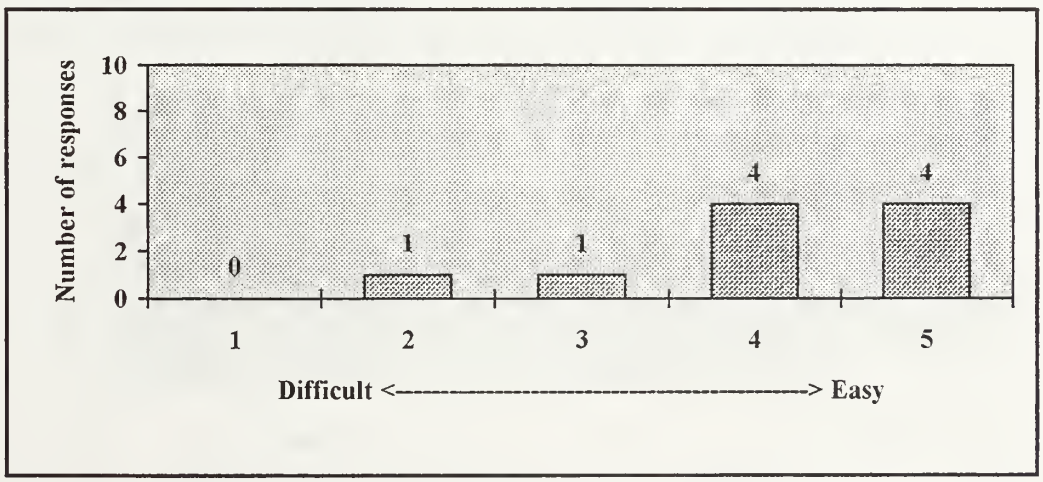


Figure 4.8. Learning to Operate the ANVIS/HUD CBT

Table 4.3. Response Statistics for Question 3.1

Statistic	Value
Mean	4.10
Median	4.00
Standard Deviation	0.994

Question 3.2: Understanding the instructions?

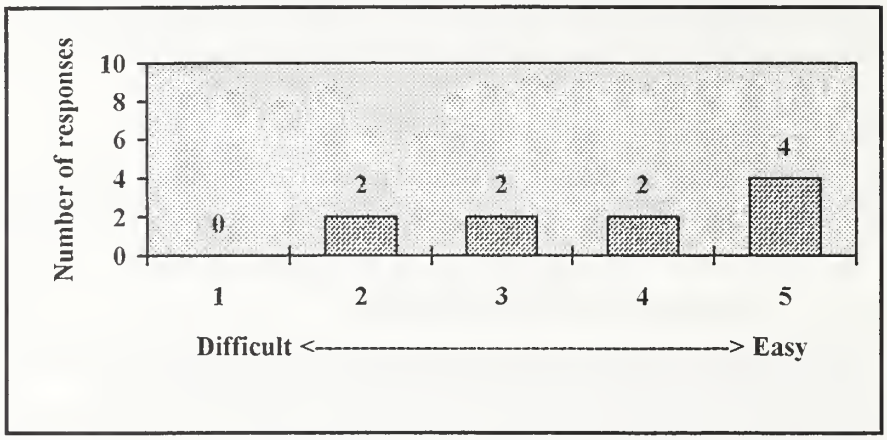


Figure 4.9. Understanding the ANVIS/HUD CBT Instructions

Table 4.4. Response Statistics for Question 3.2

Statistic	Value
Mean	3.80
Median	4.00
Standard Deviation	1.229

Question 3.3: Terminology used in the instructions?

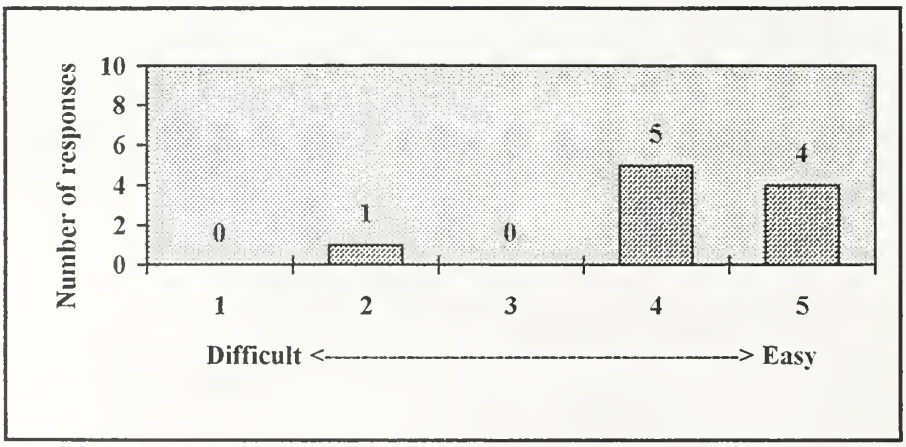


Figure 4.10. Understanding ANVIS/HUD CBT Terminology

Table 4.5. Response Statistics for Question 3.3

Statistic	Value
Mean	4.20
Median	4.00
Standard Deviation	0.919

d. Navigating through the CBT

Questions concerning navigation through the CBT received relatively less favorable responses. Navigation through the modules was not considered to be entirely clear (Figure 4.11 and Table 4.6) and the VCR type navigation controls were not easy to use for moving through lessons and modules (Figure 4.12 and Table 4.7). The next screen in a sequence was not wholly predictable (Figure 4.13 and Table 4.8) nor was going back to a previous screen felt to be easy (Figure 4.14 and Table 4.9). Determination of location within a task (beginning, middle, end) received the most unfavorable response of all questions specific to the CBT (Figure 4.15 and Table 4.10). Order of instruction and screens were both considered generally logical (Figures 4.16 and 4.17, Tables 4.11 and 4.12).

Question 4.1: Navigating through modules?

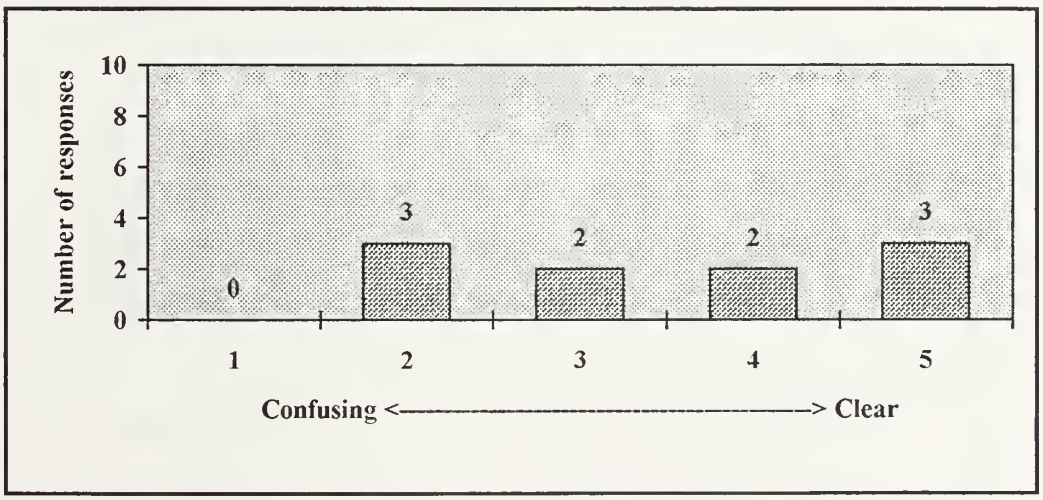


Figure 4.11. Module Navigation

Table 4.6. Response Statistics for Question 4.1

Statistic	Value
Mean	3.50
Median	3.50
Standard Deviation	1.269

Question 4.2: Use of VCR type controls for moving through lessons and modules?

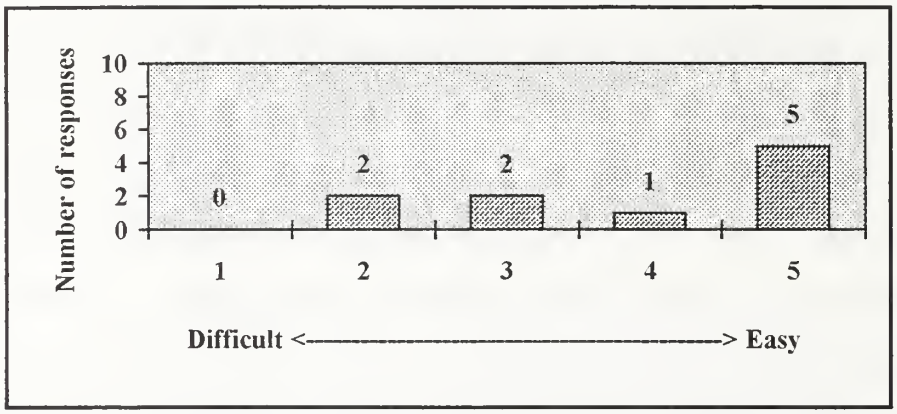


Figure 4.12. Use of VCR Type Controls

Table 4.7. Response Statistics for Question 4.2

Statistic	Value
Mean	3.90
Median	4.50
Standard Deviation	1.287

Question 4.3: Next screen in a sequence?

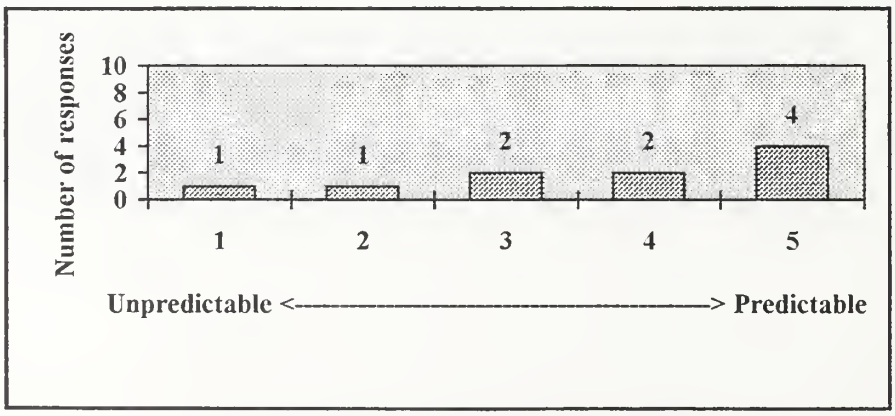


Figure 4.13. Next Screen In a Sequence

Table 4.8. Response Statistics for Question 4.3

Statistic	Value
Mean	3.70
Median	4.00
Standard Deviation	1.418

Question 4.4: Going back to previous screen?

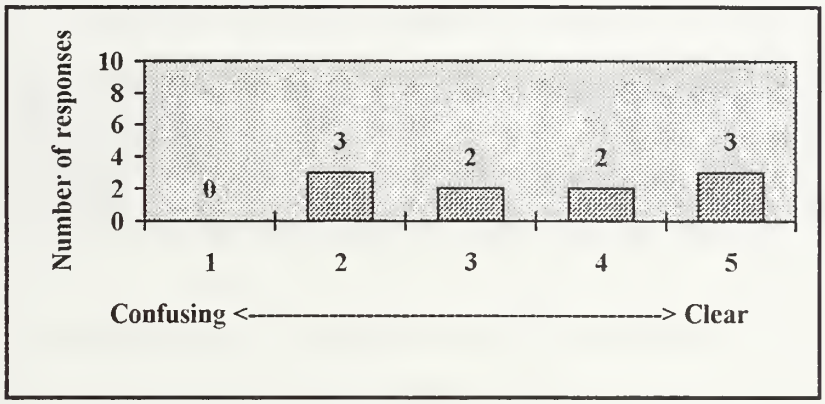


Figure 4.14. Going Back to Previous Screen

Table 4.9. Response Statistics for Question 4.4

Statistic	Value
Mean	4.00
Median	4.00
Standard Deviation	1.000

Question 4.5: Beginning, middle and end of tasks?

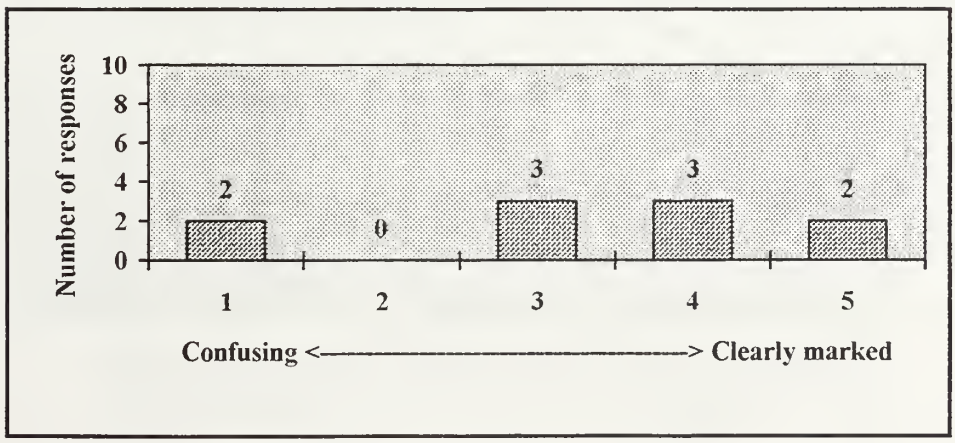


Figure 4.15. Beginning, Middle and End of Tasks

Table 4.10. Response Statistics for Question 4.5

Statistic	Value
Mean	3.30
Median	3.50
Standard Deviation	1.418

Question 4.6: Order of instruction (modules)?

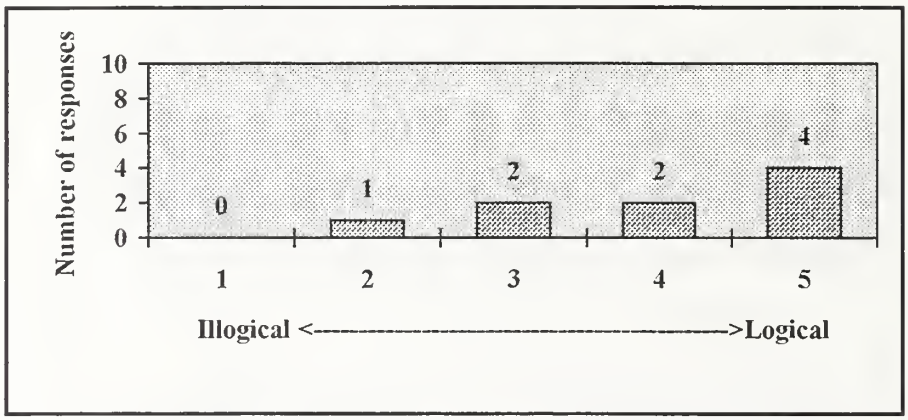


Figure 4.16. Order of Instruction

Table 4.11. Response Statistics for Question 4.6

Statistic	Value
Mean	4.00
Median	4.00
Standard Deviation	1.118

Question 4.7: Order of screens within instructional modules?

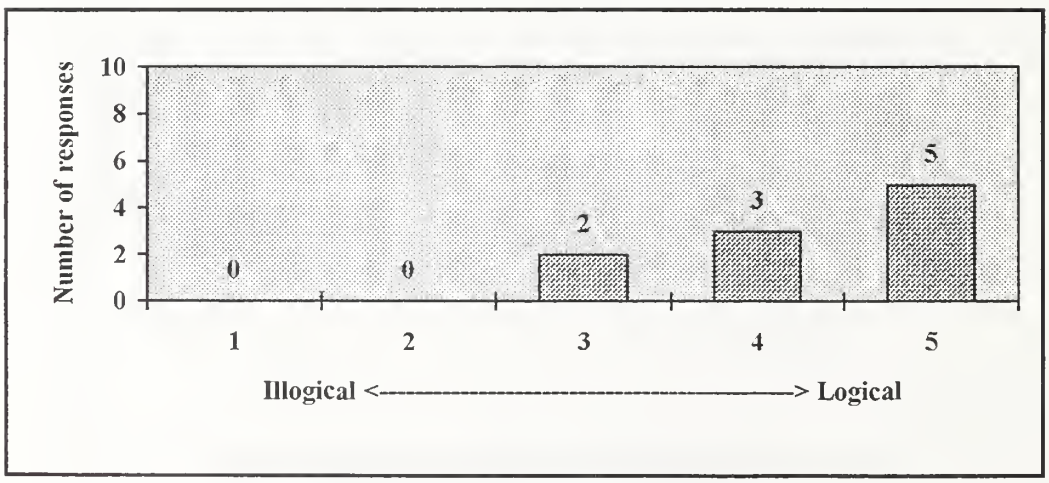


Figure 4.17. Order of Screens

Table 4.12. Response Statistics for Question 4.7

Statistic	Value
Mean	4.30
Median	4.50
Standard Deviation	0.823

e. Appearance Issues

Participants' responses to questions concerning the appearance of the CBT's interface indicated general satisfaction. Screen size was considered about right (Figure 4.18 and Table 4.13) as was size of the text or font (Figure 4.19 and Table 4.14). Graphical displays were not hard to understand (Figure 4.20 and Table 4.15), however, diagrams in the CBT were felt to be less than helpful (Figure 4.21 and Table 4.16). Overall arrangement of information on the screen was considered logical (Figure 4.22 and Table 4.17).

Question 5.1: Screen (window) size?

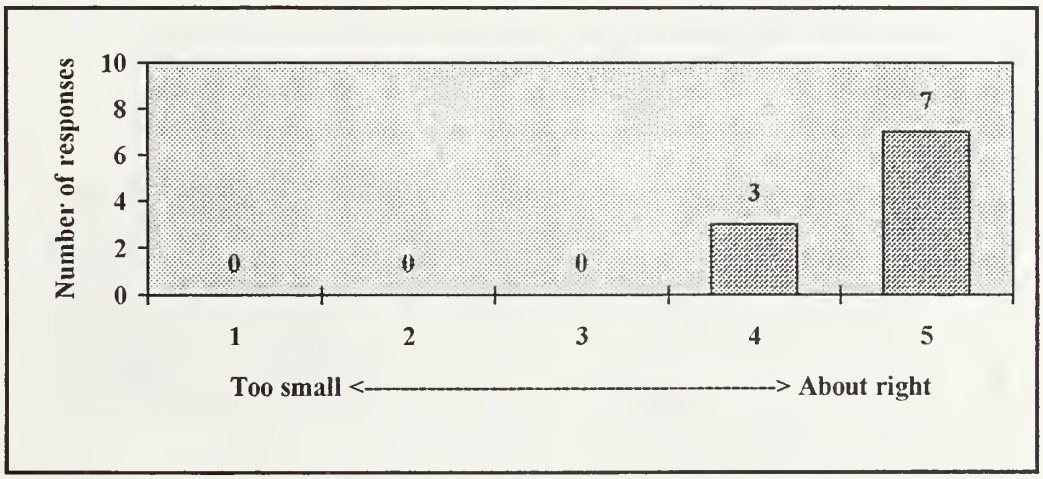


Figure 4.18. Screen Size

Table 4.13. Response Statistics for Question 5.1

Statistic	Value
Mean	4.70
Median	5.00
Standard Deviation	0.483

Question 5.2: Text (font) size?

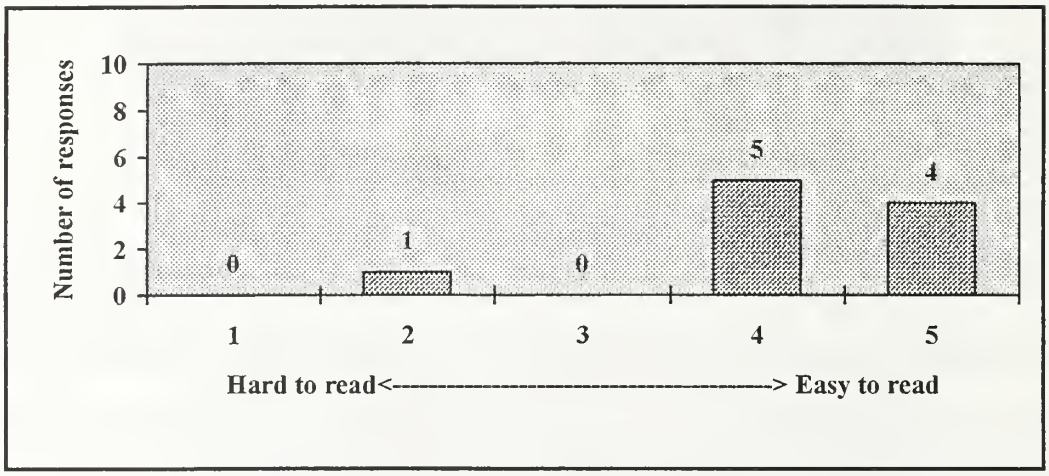


Figure 4.19. Text Size

Table 4.14. Response Statistics for Question 5.2

Statistic	Value
Mean	4.20
Median	4.00
Standard Deviation	0.919

Question 5.3: Graphical displays?

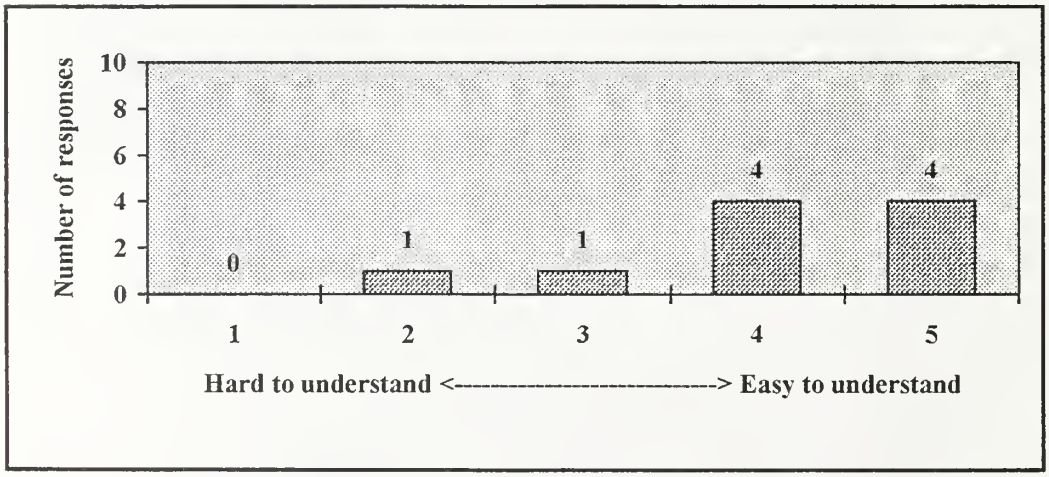


Figure 4.20. Graphical Displays

Table 4.15. Response Statistics for Question 5.3

Statistic	Value
Mean	4.10
Median	4.00
Standard Deviation	0.994

Question 5.4: Diagrams?

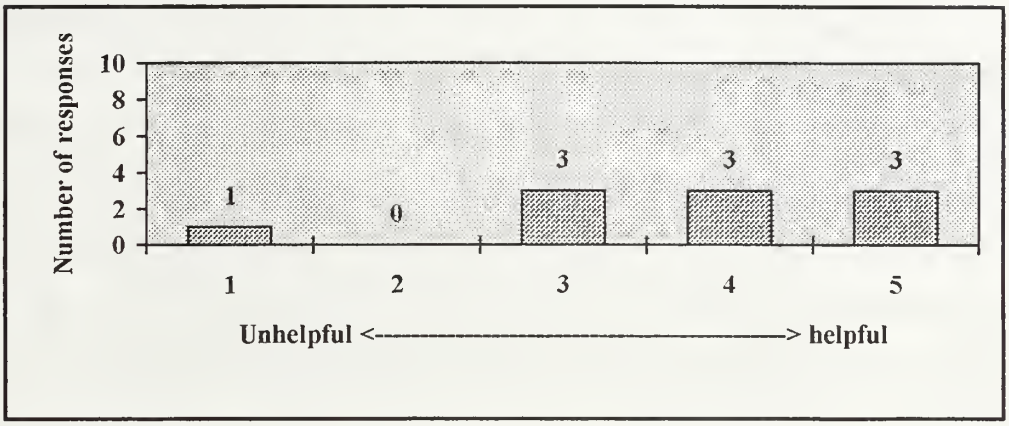


Figure 4.21. CBT Diagrams

Table 4.16. Response Statistics for Question 5.4

Statistic	Value
Mean	3.70
Median	4.00
Standard Deviation	1.252

Question 5.5 Arrangement of information on screen?

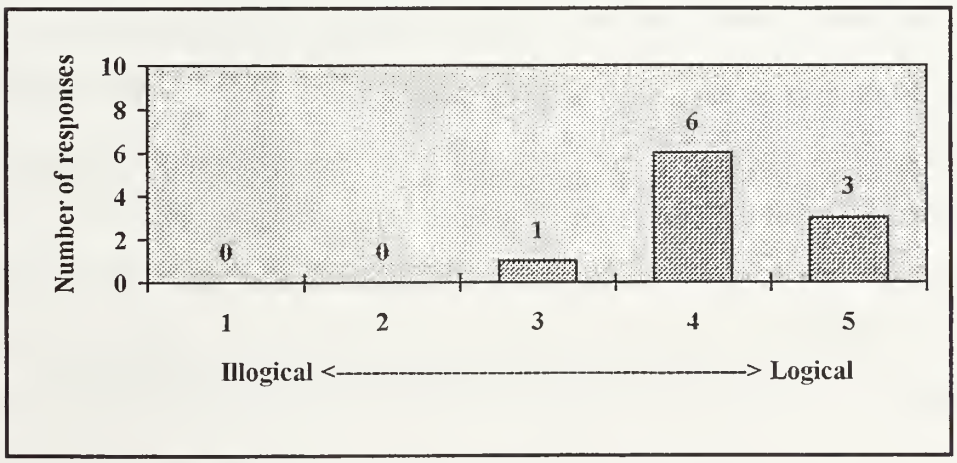


Figure 4.22. Information Arrangement

Table 4.17. Response Statistics for Question 5.5

Statistic	Value
Mean	4.20
Median	4.00
Standard Deviation	0.632

f. Time Spent Using The CBT

No participant spent more than one hour using the ANVIS/HUD CBT (Figure 4.23).

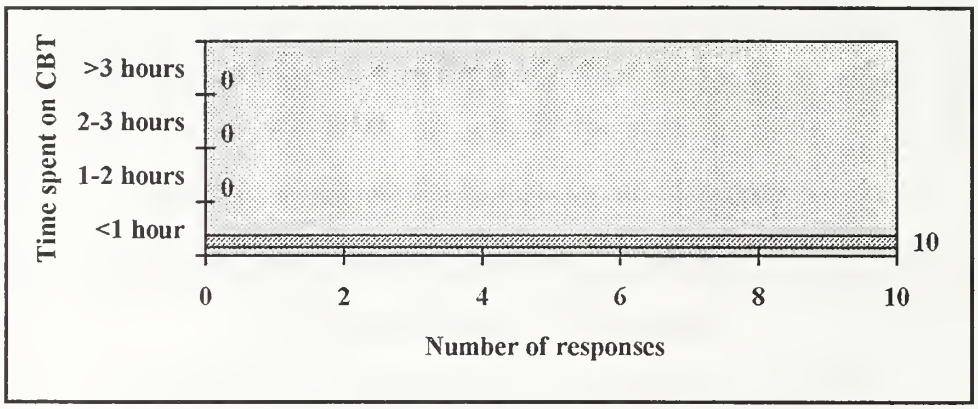


Figure 4.23. Time Spent Using CBT

g. Comments

Comments from the questionnaire (Appendix C) were generally positive about the ANVIS/HUD CBT. However, four areas were mentioned more than once and should warrant attention from a usability standpoint.

1. A desire to access definitions and explanations of acronyms from any point in a lesson.
2. A desire to have a shortened or a more generalized version of the CBT.
3. A need for indication of the beginning and end of modules and lessons.
4. A desire for a shorter, more concise introduction to the CBT.

2. Interviews

Similar to the analysis done on the questionnaire comments, interview statements were transcribed (Appendix D) and examined for multiple references to the same subject. In response to the first question concerning, the CBT as an easy to use tool for learning, one participant indicated that he had gotten lost while another said that the course flowchart in the overview was confusing.

When asked if anything was particularly difficult to do or understand, two participants expressed frustration with not knowing when the end of a lesson had been reached. Two other participants felt that the CBT introduction was difficult and time consuming.

The third question asked the participants to describe one change they would like to make the CBT. Five out of the ten participants mentioned navigation or page advancement issues. The desire for quick access to acronym definitions was expressed by two others.

The final question inquired about any features that might be added to the CBT. Again access to acronym definitions was mentioned.

3. Videotapes

Viewing the videotapes of the test sessions largely reinforces the usability issues established by the questionnaire and interviews. However, the videos did reveal other more specific usability problems.

At the completion of some lessons the user is automatically advanced to he next lesson while other lessons return the user to the first page of the completed lesson. The users then became confused as to their location within the instructional material. This occurred repeatedly, despite the presence of the location index in the lower right hand portion of the interface. Other participants became confused over the need to return to either the main menu to begin a new module or the module introduction page to begin a new lesson.

Another problem, closely related to the previously mentioned problem concerned the navigation buttons. In trying to return to previous pages, participants were observed alternately selecting the rewind and replay buttons. Neither button indicated a "not available" status and the participants became further confused when no reaction was forthcoming from the CBT.

The end-of-lesson exercises also represent a potential usability problem. The use of a green check-mark to indicate a correct answer and a red "X" as indication of an incorrect answer was not immediately obvious to many participants. Additionally some of the test participants seemed surprised at the existence of the end-of-lesson exercises.

Along with the above mentioned usability issues several minor bugs and errors were discovered during the testing and are evident on the videotapes. These bugs, while warranting attention, do not impede use of the CBT or distract significantly from the learning process.

4. Minimum System Requirements

Previously established minimum hardware requirements for operating the ANVIS/HUD were found to be inadequate during the conduct of this usability test. Those minimums called for an IBM compatible computer with a 486SX CPU operating at 33Mhz, 4 Mb RAM, 17Mb available hard disk space, VGA monitor with 256 color capable video card and Windows 3.X. However, with the ANVIS/HUD CBT installed on a machine closely matching this configuration, albeit one that was connected to a local area network (LAN), performance was degraded to the point that significant waits were encountered with all actions. This experience indicates that the current hardware requirements are probably inadequate and that testing should be done on operation of the CBT in a network environment.

C. DISCUSSION

The unfamiliarity of the participants with the ANVIS/HUD, coupled with their generally regular use of computers, make the ANVIS/HUD CBT an ideal method for instructing HH-60H pilots in the use of the ANVIS/HUD. But, as the findings above indicate, there are improvements that can be made to the ANVIS/HUD CBT that will make it a much more acceptable instructional tool.

First, the overview module which introduces users to the CBT needs to be shortened and simplified. This was emphasized by both the questionnaire and the interviews.

Second, the navigation metaphor used in the CBT presented problems on several occasions and was mentioned in every part of the usability test. The findings indicate that the VCR control interface is somewhat ambiguous. The play button suggests that the trainer might advance through the lessons automatically when "played". Additionally, the difference between the rewind and replay buttons is not intuitively obvious and tends to confuse users even though it is explained in the overview.

Third, a lack of clearly defined separations between the instructional units caused many participants to become lost in the instructional material. This is evident in the videotapes and was mentioned three times during the interviews and in the comments section of the questionnaire.

Fourth, the extensive use of acronyms throughout the instructional material was criticized repeatedly. The present structure of the CBT expects users to read the definitions during the overview before moving on to the lessons. Any need to refer to the acronym definition list requires the user to completely exit the lesson with no quick method of returning to a particular point in the lesson.

Fifth, the ANVIS/HUD CBT probably needs to be tailored to its intended user. Of particular note is the module covering maintenance procedures. Many

participants questioned the inclusion of this material and expressed a desire for a shortened and more pilot specific trainer.

Finally, and perhaps most importantly, almost all participants expressed genuine enthusiasm for the CBT and considered their use of it to be a worthwhile learning experience. Most of the problems highlighted were felt to be relatively minor by the participants and were not significant enough to cause them not to use the CBT.

V. CONCLUSION AND RECOMMENDATIONS

A. CONCLUSION

The design and development of the ANVIS/HUD CBT were done to fulfill the requirement for masters' degrees by two Naval Postgraduate school students. The product of that work, while noteworthy, is still somewhat unrefined. The ANVIS/HUD CBT offers the Navy the promise of a portable, customized instructional system that could significantly enhance the Navy's night vision training program. However, as with any software program, commercially produced or otherwise, success is determined not only by the functionality of the product, but how easy it is to use as well. Significant to ensuring ease of use is the proper application of recognized standards in human-computer interface design. Adherence to these standards during design however, does not guarantee that individuals who must use the ANVIS/HUD CBT will find it easy to do so. The results of this study serve to validate the usability testing methodology established in this study as a valuable way of ensuring usable software programs. Because usability testing bridges the gap between a finished software product and a potentially successful one its continued use by developers of future versions of the ANVIS/HUD CBT or other similar programs is strongly recommended. It is important that we do not compromise the successful adoption of new military training software by overlooking the needs of the ultimate user to have a simple to use, yet content rich tool for learning.

B. RECOMMENDATIONS

The results of this study and the conclusions drawn from them suggest that the actions listed below will enhance the ANVIS/HUD CBT and make it a much easier to use instructional tool:

- Remove the Warnings and Cautions from the overview and place them individually at the most applicable points in the lessons. This

serves to shorten the overview module and to provide the necessary context for the warnings and cautions that presently is not available.

- Move the library of terms and acronyms to a separate section and incorporate hypertext links from the instructional lessons to the individual acronyms. This allows the user to immediately access definitions and explanations of any acronyms or terms used in the lessons. This also reduces the length of the overview module.
- Add an introductory page to the beginning of each lesson. By doing so, users will immediately know that they are starting a new lesson. The lesson introduction page should briefly explain what the lesson covers, the length of the lesson and number of exercise questions at the end of the lesson.
- Add a completion page to the end of every lesson and module. Similar to the lesson introduction pages, a lesson or module completion page serves to identify the end of an instructional section and provides a timely point to end an instructional session if so desired.
- Replace the current navigation structure with a more simplified one. One alternative navigation interface could use only the following four buttons:

  1. "Next" - advances the user to the next page.
  2. "Back" - returns the user to the previous page .
  3. "Go to" - allows the user to access any section (modules, lessons or exercises) from any page in the CBT.
  4. "Cancel" - allows the user to exit the current window or the program.

- Ensure that all lessons and modules automatically advance, upon completion, to the next lesson or module. This avoids the confusion that occurs when a user is returned to the beginning of a lesson or module from the end of that lesson or module.
- Remove the module covering maintenance procedures or build a second trainer that focuses specifically on maintenance requirements. This significantly reduces the length of the CBT and places the emphasis on the knowledge required by an operator.

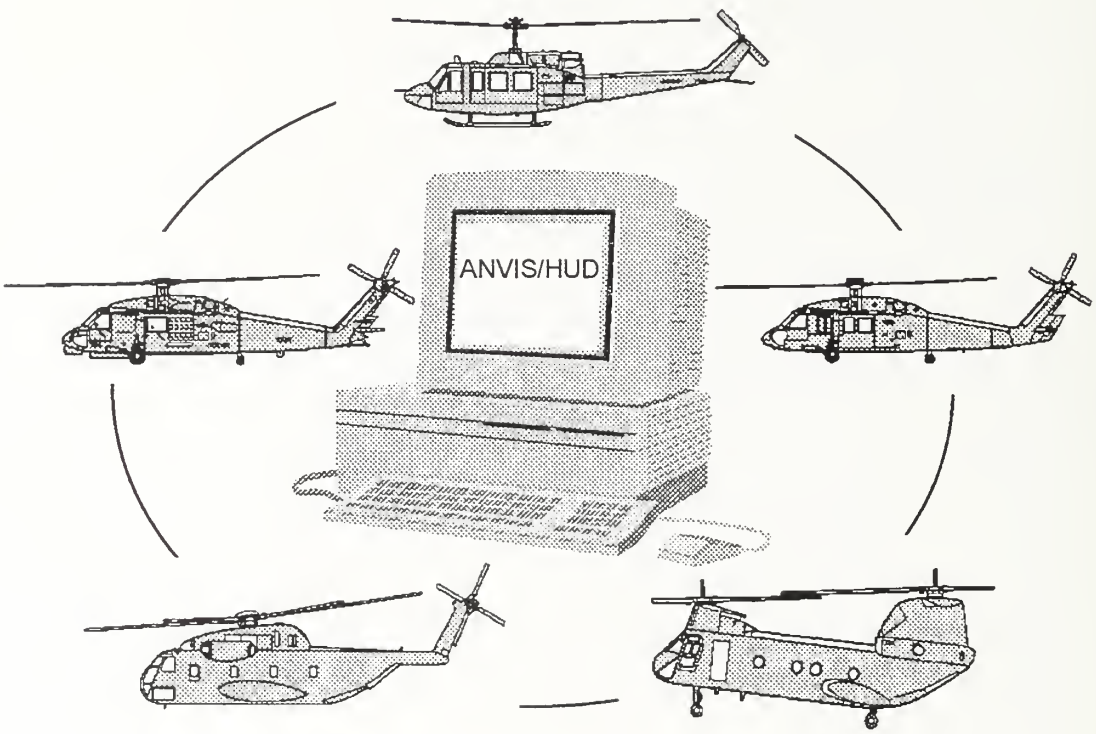
C. FUTURE RESEARCH

The recommendations offered above are an indication that future work can be done on the ANVIS/HUD CBT. Research into the construction of a hypertext linking system for the trainer is probably necessary as is a more optimum design of the navigation interface. Another area of research applicable to the ANVIS/HUD CBT concerns installation and performance of the software over a local area network (LAN). Both squadrons visited during this study were entirely networked. From a control and availability standpoint it may be better if the ANVIS/HUD CBT software could be installed on the local server and run through any of the squadrons personal computers (PC). This action provides a single point for installation of upgrades or corrections to the instructional material.

Future research involving the usability testing methodology developed by this study is also advisable. By applying the data collection methods of this study to other software programs similar to the ANVIS/HUD, a knowledge base of usability factors can be developed and used in the design of future CBT systems. The viability of establishing a usability laboratory for that purpose should also be explored.

APPENDIX A. USABILITY QUESTIONNAIRE FOR THE ANVIS/HUD CBT

Usability Questionnaire
for the
ANVIS/HUD Computer Based Trainer



Questionnaire for ANVIS/HUD Computer Based Trainer Users

The ANVIS/HUD Computer Based Training software was developed at the Navy's School of Aviation Safety specifically for the HCS squadrons. By completing this questionnaire you will greatly assist us in improving this product. The first few questions below cover your knowledge of the ANVIS/HUD and computers in general. The remaining questions deal exclusively with the trainer software and will help us to tailor it for your ease of use and learning. Finally, space is provided at the end of this questionnaire for your comments concerning the ANVIS/HUD CBT. Your time spent on this questionnaire is appreciated.

Please circle the number or phrase that best reflects your answer to the question.

1. Background

1.1 How familiar are you with the ANVIS/HUD?

Not at all familiar Very Familiar
1 2 3 4 5

1.2 Have you flown with the ANVIS/HUD?

No Yes

2. Computer Usage

2.1 Are you familiar with the "Windows" type operating system?

Not at all familiar Very Familiar
1 2 3 4 5

2.2 How often do you use a computer?

Rarely Occasionally Frequently Constantly

2.3 Do you own a home computer?

No Yes

2.4 What activities do you most often use the computer for? (Circle all that apply)

Word processing Spreadsheets
Games Database
Graphics Education
Online/Internet

2.5 How would you describe yourself with respect to computers?

Non-user Novice Regular user Hobbyist A real geek

The following questions apply to your experience with the Computer based training software for the ANVIS/HUD. Please circle the number on the rating scale that most closely reflects your attitude or opinion on the statements under each section below. If you have no opinion or feel the statement does not apply to your experience with the trainer please circle "NA".

3. Getting Started

3.1 Learning to operate the program	difficult 1 2 3 4 5 easy	NA
3.2 Understanding the instructions	difficult 1 2 3 4 5 easy	NA
3.3 Terminology used in the instructions	confusing 1 2 3 4 5 clear	NA

4. Navigation

4.1 Navigating through modules	confusing 1 2 3 4 5 clear	NA
4.2 Use of "VCR" type controls for moving through lessons and modules	difficult 1 2 3 4 5 easy	NA
4.3 Next screen in a sequence	unpredictable 1 2 3 4 5 predictable	NA
4.4 Going back to previous screen	impossible 1 2 3 4 5 easy	NA
4.5 Beginning, middle and end of tasks	confusing 1 2 3 4 5 clearly marked	NA
4.6 Order of instruction (modules)	illogical 1 2 3 4 5 logical	NA
4.7 Order of screens within instructional modules	illogical 1 2 3 4 5 logical	NA

5. **Appearance**

5.1 Screen (window) size	too small				about right	
	1	2	3	4	5	NA
5.2 text size (font)	hard to read				easy to read	
	1	2	3	4	5	NA
5.3 Graphical displays	hard to understand				easy to understand	
	1	2	3	4	5	NA
5.4 Diagrams	unhelpful				helpful	
	1	2	3	4	5	NA
5.6 Arrangement of information on screen	illogical				logical	
	1	2	3	4	5	NA

6. How much time did you spend using the CBT?

Less than 1 hour 1-2 hours 2-3 hours 3 hours or more

END

Comments: (Tell us what you think of the ANVIS/HUD trainer software.)

Copyright © 1989, 1994
University of Maryland at College Park
All Rights Reserved

APPENDIX B. RESEARCH PARTICIPANT CONSENT FORM

ANVIS/HUD CBT SOFTWARE USABILITY STUDY

CONSENT TO BE A RESEARCH PARTICIPANT

A. PURPOSE AND BACKGROUND

LCDR Dan Rozelle, of the Naval Postgraduate School, is doing a study on the usability of the ANVIS/HUD Computer-Based Trainer (CBT) software. I have been asked to participate in this study.

B. PROCEDURES

If I agree to be in the study, the following will occur:

1. I will use the ANVIS/HUD CBT software.

2. If I agree, I will be videotaped using the ANVIS/HUD CBT software.

3. I will be asked to respond to a questionnaire and to answer questions about my experience with the ANVIS/HUD CBT software.

These procedures will be done at my place of work (HCS Squadron) and will take a total time of about 2 hours.

C. RISKS/DISCOMFORTS

1. If, at any time, I feel uncomfortable with the procedures of this study I will be able to stop at any time.

2. I am free to decline to answer any questions I do not wish to, or to stop the interview at any time.

D. BENEFITS

There will be no direct benefit to me from participating in this study. The anticipated benefit of these procedures is a better understanding of improvements that may be made to enhance the usability of the ANVIS/HUD CBT software.

E. CONFIDENTIALITY

Study records will be kept as confidential as is possible. No individual identities will be used in any reports or publications resulting from the study. Study information will be coded and held by the School of Aviation Safety at the Naval Postgraduate School. Only study personnel will have access to the files and videotapes.

F. ALTERNATIVES

I am free to choose not to participate in this study.

G. COSTS

There will be no costs to me as a result of taking part in this study.

H. REIMBURSEMENT

I will not be reimbursed for my participation in this study.

I. QUESTIONS

I have talked to LCDR Rozelle about this study, and have had my questions answered. If I have any further questions about this study, I may call him at (408) 372-2184. If I have questions specific to the ANVIS/HUD CBT software they should be directed to Dr. Anthony Ciavarelli of the School of Aviation Safety at the Naval Postgraduate School. I can contact him at DSN 878-2581 or Commercial (408) 656-2581.

J. CONSENT

I have been given a copy of this consent form to keep.

PARTICIPATION IN RESEARCH IS VOLUNTARY. I am free to decline to be in this study. If I elect to participate, I may withdraw from the study at any point.

_____ _____

Date Subject's Signature

_____

Person Obtaining Consent

APPENDIX C. COMMENTS TAKEN FROM QUESTIONNAIRE

- Excellent program. Easy to understand and a great initial introduction to the system
- Overall, very informative for basic use, setup and terminology. Excellent for supplementing the NATOPS manual.
- The only changes I would suggest are:
  1) Change font of letters (larger).
  2) The maintenance page/procedure page needs to be all one color (yellow or white).
  3) Describe what the thrust lever is.
  4) Have an icon that allows you to go directly to the acronyms definition page from any lesson.
- My only two comments concern knowing when you reach the end of a module. An indication that you have completed a module and need to return to the main menu would be an asset. Also, hyperlinks to acronyms and definitions would make the system easier to use.
- Great. Ensure the setup instructions include recommendations to use a 486/66 or higher machine and unhook machine from any LAN it is on. (Note: This particular comment was directed toward slowness of the system experienced at HCS-4. This is addressed in the Other Findings sections.)
- Overall - good. Intro is too long and over explains many obvious processes. It's better to learn as you go. Unable to confirm when you have completed a module. Too easy to get lost. Need feedback on whether a key [button] has been selected or is available is inadequate.
- I feel the trainer would be an invaluable tool prior to using the ANVIS/HUD for the first time, and then as a valuable training aid afterwards. The best possible situation would be to somehow have a mock collective stick connected to an ANVIS/HUD and a PC to provide the most realistic training prior to flying the HUD in the aircraft. Initially, I believe the HUD will be a distraction from flying the aircraft. It will be essential to have a syllabus designed using this type of computer simulation.

- There needs to be a shortened version of this course that tells a pilot the essentials of what he needs to know about the control panel. Pilots can figure out how to mount the ANVIS/HUD and plug it in with one demo (a Fam 0). After that all a pilot wants to know is what does the control box do and how can it be operated in flight. You could still have the long version of the course for the detailed info concerning the entire system.
- The beginning tutorial needs help. Look at the current Microsoft Word tutorial as an example. Verbiage should be clear and concise, with animation to back it up rather than the long winded college level verbosity.
- Should be a good intro for first timers. [You] might consider a second unit for refresher training to avoid rehashing something not needed the second time.

APPENDIX D. INTERVIEW QUESTIONS AND SIGNIFICANT RESPONSES

Question: Did you find the portion of the CBT that you worked with to be an easy to use tool for learning?

Responses (HCS-4):

- Yes, but there should be some indication that there will be exercises at the end of the lessons. If I would have known I wouldn't have gone through the lessons quite so fast
- Actually yes. Before using the HUD for the first time it would be essential.
- Yes. The course overview had a flowchart that was initially confusing but I thought that overall it was pretty good. The one module that I worked through, I thought was fairly well laid out.
- Not completely. Generally yes, but I got a little lost in where I was. Although there is a page number, unless you are really paying attention to where you are, you can easily get lost or skip a page.

Responses (HCS-5):

- Yes, I thought it was great
- No, it was difficult to use

Question: Was there anything that you found particularly difficult to do or understand?

Responses (HCS-4):

- No, except for getting lost in some of the menus and not knowing where the lessons ended.
- Not really, but I didn't think that all the maintenance stuff was necessary for the pilots
- The acronyms, and also I wasn't sure whether I'd actually reached the end of a lesson. Maybe an end of lesson indicator would be helpful.

- The introduction was tedious and over-explained. In some of the sentences the English was weird. If you read on, it makes sense but if you don't, it doesn't make sense.
- The rewind and replay button. The rewind didn't seem to work part of the time. And …you can't tell if you are doing an illegal request. They [the buttons] could change color or say "not available" so that they are not selectable.

Responses (HCS-5):

- The electrical diagram was confusing, it should include more about what each electrical module does.
- The course introduction was painful

Question: If you could change one thing about the CBT to make it easier to use, what would it be?

Responses (HCS-4):

- The control buttons. The current setup is kind of confusing.
- Need to clear up the navigation problems so that it's easier to navigate through.
- If, when you answered a question, whether right or wrong, it would automatically go to the next question instead of having to hit the "next question" button. I just thought that it was time consuming.
- Maybe hyperlinks for the acronyms.
- I don't think there is anything that needs to be changed. There is the matter of figuring out the setup that is being used. This one [CBT] is setup like a tape recorder which isn't hard to figure out.
- Simply making this [the navigation panel] easier to use. Also is there a way to tell if you've done every page. There is no way that I can tell if I missed a page.

Responses (HCS-5):

- Lock the cursor (mouse arrow) to automatically go the default selection button.
- Replace the introduction with a more state-of-the-art tutorial.

- When an acronym is used in the text it should be spelled out immediately prior. Continuous repetition of the acronym and what they mean acts as a learning tool. A "Click-on" feature (hypertext) in the text should enable a detailed explanation or diagram of the selected acronym (component).

Question: Are there any features you would like to see added to the CBT?

Responses (HCS-4):

- It would be nice to have some indication that the next page is loading. (Note: this particular response is probably due to the slow speed of the computer being used.)
- I only used the "Play" button at the bottom. Do all the other buttons do something?
- Definitely the acronym thing. If there is an acronym I want to click on it to see what it means.

Responses (HCS-5):

- Animation.

APPENDIX E. SQUADRON PRESENTATION

LCDR Dan Rozelle

Information Technology Management
Curriculum
at the
Naval Postgraduate School

Improving the Usability of Helicopter ANVIS/HUD Computer-Based Training Systems

...to gather your reactions and opinions on the ANVIS/HUD CBT

So what is the ANVIS/HUD CBT?

Stand-alone Computer-Based Trainer that teaches ANVIS/HUD symbology, basic system operating procedures, and maintenance.

- Designed to run on any Windows based PC

Background

- In 1992 a comprehensive study of NVG training was conducted
- Participants included NPS, Aviation Safety School, NAVAIR, Naval Systems Training Center
- Results showed the need for improved SOPs, revision of training doctrine and enhanced training equipment

Background

- Multimedia Development Lab established at NPS under direction of Dr. Tony Ciavarelli
- Several theses completed researching improved NVG training

ANVIS/HUD CBT

- Developed by Shaffer and Kern in 1996
- Based on system originally designed for UH-1
- Consists of 4 training modules and an Overview

ANVIS/HUD CBT
Overview

- Describes Structure of Trainer
- Demonstrates Navigation of Trainer
- Library of Terms, Cautions and Warnings

ANVIS/HUD CBT
Module 1

HUD System Overview

Lesson 1 - Introduction to the ANVIS/HUD System

Lesson 2 - Basic ANVIS/HUD Components

ANVIS/HUD CBT
Module 2

HUD Symbology Set

Lesson 1 - HUD Symbology Selection

Lesson 2 - HUD Pitch and Roll Adjustment

ANVIS/HUD CBT
Module 3

HUD Operating Procedures

Lesson 1 - Installing Display Equipment
Lesson 2 - Starting the HUD System
Lesson 3 - Using the HUD System
Lesson 4 - Egress Procedures

ANVIS/HUD CBT
Module 4

HUD Operating Procedures

Lesson 1 - HUD O-Level Maintenance

Lesson 2 - Electro-Static-Discharge Considerations

Lesson 3 - HUD Unit Descriptions and Maintenance Procedures

ANVIS/HUD CBT
Module 4 (cont.)

HUD Operating Procedures

Lesson 4 - HUD Receipt, Shipment, and Storage Procedures

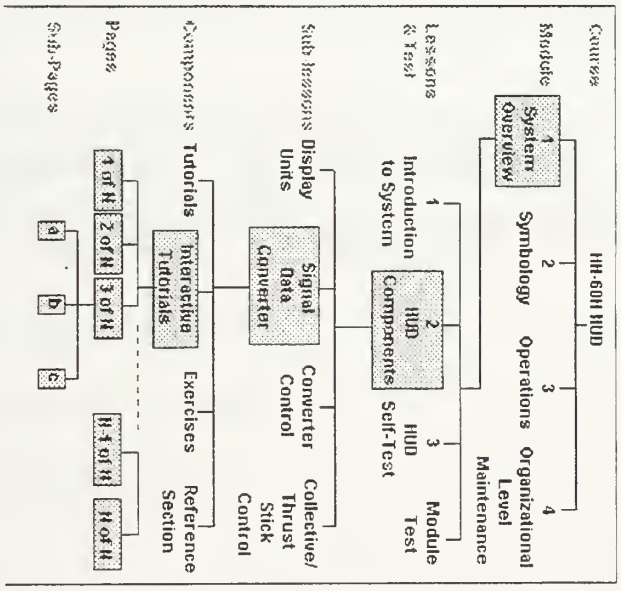
Lesson 5 - HUD Operational Checks

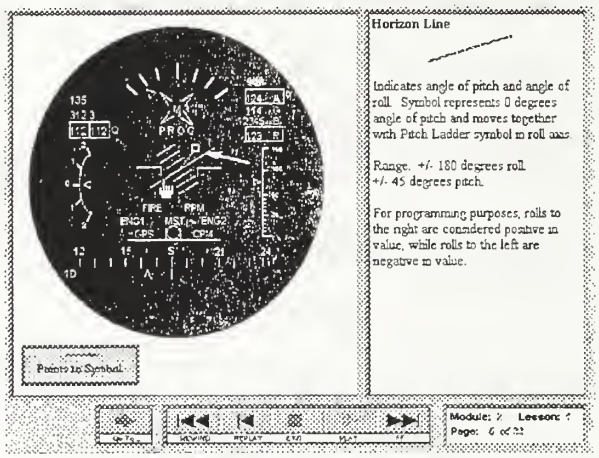
Lesson 6 -Fault Isolation and Troubleshooting

ANVIS/HUD CBT
Modules Under Development

Module 5 - Human Factors

Module 6 - Basic ANVIS/HUD Overview




Horizon Line
Indicates angle of pitch and angle of roll. Symbol represents 0 degrees angle of pitch and moves together with Pitch Ladder symbol in roll axis.
Range: +/- 180 degrees roll
+/- 45 degrees pitch.
For programming purposes, rolls to the right are considered positive in value, while rolls to the left are negative in value.
Points to Symbol
PROG
FIRE
PPM
Module: 2 Lesson: 1
Page: 6 of 22

...to gathering your reactions and opinions on the ANVIS/HUD CBT

Usability Testing

- Done by most major software vendors
- Focus is on **usability** vice **functionality**
- Attempts to answer questions such as...

- Are users able to quickly learn how to use the program?
- Is navigation within the program easy to accomplish?
- Is the text easy to read and understand?
- Graphics and diagrams?
- Arrangement of information?
- Sequence of screens?

Usability Testing

- Volunteers are needed to…
  - use the CBT (Overview and Module 1) while being videotaped
  - fill out a brief questionnaire
  - answer a few short questions

Questions?

APPENDIX F. SPREADSHEET DATA FOR THE ANVIS/HUD CBT USABILITY TEST QUESTIONNAIRE

	Q 1.1	Q 1.2	Q 2.1	Q 2.2	Q 2.3	Q 2.4	Q 2.5	Q 3.1	Q 3.2	Q 3.3
Respondent 1	1	No	5	Constantly	Yes	wp	Regular user	5	5	5
Respondent 2	2	No	4	Frequently	Yes	wp, s, d	Regular user	4	4	4
Respondent 3	3	No	4	Constantly	Yes	wp, o/I, s, e,	Regular user	4	4	4
Respondent 4	3	No	3	Occasionally	Yes	wp, g, d	Novice	5	5	5
Respondent 5	3	No	5	Frequently	Yes	wp, g, o/I,	Regular user	5	5	4
Respondent 6	2	No	4	Constantly	Yes	wp, o/I, s, d	Regular user	3	2	4
Respondent 7	2	No	4	Frequently	Yes	wp, e	Regular user	4	3	4
Respondent 8	3	No	4	Frequently	Yes	wp, g, o/I, d	Regular user	4	3	5
Respondent 9	2	No	5	Constantly	Yes	wp, gr, o/I, s, d, e	A real geek	2	2	2
Respondent 10	3	No	5	Constantly	Yes	wp, o/I, d, e	Regular user	5	5	5
Mean	2.40		4.30					4.10	3.80	4.20
Median	2.50		4.00					4.00	4.00	4.00
Standard Deviation	0.699206		0.674949					0.994429	1.229273	0.918937

	Q 4.1	Q 4.2	Q 4.3	Q 4.4	Q 4.5	Q 4.6	Q 4.7	Q 5.1	Q 5.2	Q 5.3	Q 5.4
Respondent 1	5	5	5	5	5	5	5	5	5	5	3
Respondent 2	5	5	4	4	4	5	5	5	4	4	4
Respondent 3	3	3	3	3	3	4	4	4	2	3	3
Respondent 4	3	5	5	4	3	5	5	5	4	5	5
Respondent 5	4	5	5	5	5	5	5	5	5	5	5
Respondent 6	2	3	4	2	4	4	4	5	4	5	5
Respondent 7	4	2	2	NA	3	2	3	4	4	4	4
Respondent 8	2	2	1	4	1	3	4	4	4	4	4
Respondent 9	2	4	3	4	1	3	3	5	5	2	1
Respondent 10	5	5	5	5	4	NA	5	5	5	4	3
Mean	3.50	3.90	3.70	4.00	3.30	4.00	4.30	4.70	4.20	4.10	3.70
Median	3.50	4.50	4.00	4.00	3.50	4.00	4.50	5.00	4.00	4.00	4.00
Standard Deviation	1.269296	1.286684	1.418136	1	1.418136	1.118034	0.823273	0.483046	0.918937	0.994429	1.251666

	Q 5.5	Q 6
Respondent 1	5	<1 hour
Respondent 2	4	<1 hour
Respondent 3	3	<1 hour
Respondent 4	5	<1 hour
Respondent 5	5	<1 hour
Respondent 6	4	<1 hour
Respondent 7	4	<1 hour
Respondent 8	4	<1 hour
Respondent 9	4	<1 hour
Respondent 10	4	<1 hour

Mean	4.20
Median	4.00
Standard Deviation	0.632456

LIST OF REFERENCES

Brown, C., *Human-Computer Interface Design Guidelines*, Ablex Publishing Corporation, 1988.

Bryant, Bobby and Glenroy Day Jr., *A Computer-based Multimedia Prototype For Night Vision Goggles*, Master's Thesis, Naval Postgraduate School, Monterey, California, September 1994.

Chin, John P., Virginia A. Diehl and Kent L. Norman, "Development of an Instrument Measuring User Satisfaction of the Human-Computer Interface," *Proceedings of CHI, 1988* (Washington D.C., May 15-19, 1988), pp. 213-218, ACM, New York, 1988.

Ciavarelli, Anthony and Kishore Sengupta, *Night Vision Goggle (NVG) Training Technology Report*, Technology Report, Naval Postgraduate School, Monterey, California, December, 1994.

Desurvire, H., J. Kondziela and M. Atwood, "What Is Gained and Lost When Using Evaluation Methods Other Than Empirical Testing," *Proceedings of CHI, 1992* (Monterey, California, May 3-7, 1992), pp. 1-16, ACM, New York, 1992.

Dumas, Joseph S. and Jancie C. Redish, *A Practical Guide to Usability Testing*, Ablex Publishing Corporation, 1993.

Epperson, Sean T., *Animation Within A Multimedia Training System For Night Vision Goggles*, Master's Thesis, Naval Postgraduate School, Monterey, California, March, 1995.

Jeffries, Robin, James R. Miller, Cathleen Wharton and Kathy M. Uyeda, "User Interface Evaluation in the Real World: A Comparison of Four Techniques," *Proceedings of CHI, 1991* (New Orleans, Louisiana, April 28-May 2, 1991), pp. 119-124, ACM, New York, 1991.

Karat, Clare-Marie, Robert Campbell and Tarra Fiegel, "Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation," *Proceedings of CHI, 1992* (Monterey, California, May 3-7, 1992), pp. 397-404, ACM, New York, 1992.

Lewis, C., P. Polson, C. Wharton and J. Rieman, "Testing a Walkthrough Methodology for Theory-Based Design of Walk-Up-and-Use Interfaces," *Proceedings of CHI, 1990* (Seattle, Washington, April 1-5, 1990), pp. 235-242, ACM, New York, 1990.

Mayhew, D., *Principles and Guidelines in Software User Interface Design*, Prentice Hall, 1992.

Meza, Fransisco Q., *Developing Multimedia Instructional Systems: An Example Application For Training In Night Vision Goggles*, Master's Thesis, Naval Postgraduate School, Monterey California, March, 1995.

Microsoft Corporation, Microsoft Usability Group Homepage, [Online] Available http://www.microsoft.com/usability/, January 25, 1997.

Nielsen, Jakob, and Rolf Molich, "Heuristic Evaluation of User Interfaces," *Proceedings of CHI, 1990* (Seattle, Washington, April 1-5, 1990), pp. 249-256, ACM, New York, 1990.

Nielsen, Jakob, "Finding Usability Problems Through Heuristic Evaluation," *Proceedings of CHI, 1992* (Monterey, California, May 3-7, 1992), pp. 373-380, ACM, New York, 1992.

Nielsen, Jakob, *Usability Engineering*, Academic Press, 1993.

Nielsen, Jakob, and Robert Mack, *Usability Inspection Methods*, John Wiley & Sons, 1994.

Nielsen, Jakob, *Severity Ratings for Usability Problems*, [Online] Available http://www.useit.com/papers/heuristic/severityrating.html, February 1996.

Rowley, David E., "Usability Testing in the Field: Bringing the Laboratory to the User," *Proceedings of CHI, 1994* (Boston, Massachusetts, April 24-28, 1994), pp. 252-257, ACM, New York, 1994.

Shaffer, Kenan J., and Dabney R. Kern, *ANVIS/HUD Computer-Based Training System for the HH-60H Helicopter*, Master's Thesis, Naval Postgraduate School, Monterey California, May, 1996.

Shneiderman, Ben, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, Addison-Wesley, 1992.

Smith, S. and J. Mosier, *Design Guidelines for Designing User Interface Software*, Technical Report MTR-10090, The MITRE Corporation, Bedford, Massachusetts, 1986.

Wharton, C., J. Bradford, R. Jeffries and M. Franzke, "Applying Cognitive Walkthroughs to More Complex User Interfaces: Experiences, Issues, and Recommendations," *Proceedings of CHI, 1992* (Monterey, California, May 3-7, 1992), pp. 381-388, ACM, New York, 1992.

INITIAL DISTRIBUTION LIST

1. Defense Technical Informational Center 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, VA 22060-6218

2. Dudley Knox Library 2
   Naval Postgraduate School
   411 Dyer Rd.
   Monterey, CA 93943-5101

3. Anthony Ciavarelli, Code 10 2
   Aviation Safety Program
   Naval Postgraduate School
   Monterey, CA 93943-5101

4. William J. Haga, Code SM/Hg 1
   Department of Systems Management
   Naval Postgraduate School
   Monterey, CA 93943-5101

5. Commander 1
   Naval Air Systems Command
   (PMA-205) Training Systems
   1421 Jefferson Highway (JP-1, Rm 306)
   Arlington, VA 22243-1205

6. LCDR Dan Rozelle 1
   495 Mar Vista Drive
   Monterey, CA 93940

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY CA 93943-5101

DUDLEY KNOX LIBRARY
3 2768 00339217 6